Risk neutral versus real-world distribution on puclicly listed bank corporations

Michel Dacorogna, Juan-José Francisco Miguelez, Marie Kratz

▶ **To cite this version:**

Michel Dacorogna, Juan-José Francisco Miguelez, Marie Kratz. Risk neutral versus real-world distribution on puclicly listed bank corporations. 2016. hal-01373071

HAL Id: hal-01373071

https://essec.hal.science/hal-01373071

Preprint submitted on 28 Sep 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RISK NEUTRAL VERSUS REAL-WORLD DISTRIBUTION OF PUBLICLY LISTED BANK CORPORATIONS

RESEARCH CENTER

MICHEL DACOROGNA, JUAN-JOSÉ FRANCISCO MIGUELEZ, MARIE KRATZ

ESSEC WORKING PAPER 1614

JULY 2016

Risk Neutral versus Real-World Distribution of Publicly Listed Bank Corporations[1]

Michel Dacorogna (SCOR), Juan-José Francisco Miguelez (ESSEC),

and Marie Kratz (ESSEC CREAR)

ABSTRACT: in this study, we examine different quantitative methods to recover the risk neutral distribution function associated to the prices of option on bank shares. This is useful for a wide range of applications, such as determining the implicit State guarantee that systemic financial institutions benefit from the State, or looking if the market prices correctly the fat tails of financial returns. We assess the performance of these techniques in various ways, including comparing market option prices and historical Values-at-Risk to option prices and Value-at-Risk implied by the estimated risk neutral distribution. We find that, contrary to what is expected for a market composed of risk averse investors, the latter is much smaller than the one obtained from real data. We discuss our results with respect to the theory of risk neutral valuation and investor risk preference.

Keywords: *extremes; fat tail; option pricing; real world probability; risk neutral probability; SIFI; value-at-risk*

[1] This empirical study has been performed by Juan-José Francisco, ESSEC student in *M.Sc. in Management (Grande École)* and in the ESSEC-ISUP actuarial track, under the supervision of Dr. M. Dacorogna (SCOR) and Prof. M. Kratz (ESSEC, CREAR)

Table of contents

I – Introduction

I.1/ Background

Since the seminal paper of Black, Scholes and Merton in 1973 [1], the theory of derivative pricing has become a rich and profuse field of finance. Its influence goes beyond the realm of academic research, as it has also had a profound and lasting impact in trading floors all over the globe: for example, nowadays traders quote option prices on an implied volatility basis, which can be calculated through the Black-Scholes formula.

One of the fundamental notions of derivative pricing is risk neutrality. Risk neutrality implies that an agent is indifferent to risk. To illustrate this concept, suppose an economic agent is given the following choice:

- He can pocket immediately €5;
- He can toss a fair coin and get €10 if he gets heads and nothing if he gets tails.

In both cases, the expected profit is equal to €5. If the agent prefers the first option, we say he is risk averse; if he would rather select the second option, he is said to be risk prone; finally, if he is indifferent to either option, the agent is neutral towards risk.

Financial theorists have showed that, under the Absence of Arbitrage Opportunities (AOA) assumption – which postulates that an economic agent cannot generate a sure profit without taking any risk – the price of a contingent claim – or derivative – can be easily obtained by assuming that economic agents are neutral to risk. It is very important to understand that this assumption does not refer to the real world behavior of investors: it is a mathematical trick that allows generalizing and streamlining the pricing process.

This property is extremely powerful and useful, as it enables to simplify the process of pricing derivatives. For example, financial theory has traditionally worked with discount factors that should reflect an asset's underlying risk: if two securities x_1 and x_2 generate payoffs p_1 and p_2 respectively, such that x_1 is riskier than x_2, then p_1 should be discounted at a higher rate than p_2. The issue of this approach is that estimating discount factors for p_1 and p_2 is not at all a trivial problem. However, by introducing the notion of risk neutrality, the theory allows to take the risk-free rate r as the discount rate applicable to every security. Determining a suitable risk-free rate r is a much more tractable problem that estimating a discount rate μ_x for each security.

Moreover, from a mathematical perspective the notion of risk neutrality pertains to probability theory: the price of a contingent claim is given by its expected payoffs with respect to the risk neutral distribution discounted by the risk-free rate. As a consequence, the development of risk neutral (RN) pricing in the last 25 years has encouraged academics and practitioners to study and analyse the probabilistic information that can be extracted from financial securities prices. Some examples of this trend are the following:

- In the Fed Funds futures market, traders extract probabilities for Fed monetary policy decisions from the prices at which futures are quoted [13];
- In the Credit Default Swap (CDS) market, it is possible to obtain probabilities of default for the company the CDS is written on [21].

This report is in keeping with this philosophy.

Our original goal was to price the implicit protection from states that global financial institutions enjoy. During the 2008-2009 financial crisis, banking corporations which failure could have triggered a collapse of financial markets were recapitalized by their respective governments – *e.g.* Citigroup in the US or UBS in Switzerland. As a response to the threat big global financial institutions pose to financial markets, the Financial Stability Board (FSB) and the Basel Committee on Banking Supervision (BCBS) have since introduced the concept of Systematically Important Financial Institutions (SIFI) and tailored specific capital and regulatory requirements targeted to them [12].

However, by formally identifying the institutions that policy makers think can short-circuit the normal functioning of markets in the event of a crisis, regulators might actually have reinforced the implicit protection these institutions benefit from: by being forced to recapitalize ailing SIFIs if they want to maintain financial stability, as they have done extensively during the crisis, governments implicitly insure SIFIs stockholders and bondholders against bankruptcy, therefore reducing the SIFI's cost of capital. As of today, SIFIs do not have to pay for this implicit protection.

By modelling this guarantee as a put option written on the bank's assets, as suggested by M. Dacorogna [3,4], it is possible to leverage the risk neutral valuation theory to price SIFIs implicit insurance against bankruptcy and its impact on capital cost. This can be done by extracting probabilistic information – more precisely, the RN cumulative distribution function and the RN density – from option prices written on banks stocks.

However, during our research, we concluded that our technique might yield an inaccurate price. Indeed, by extracting and analysing the risk neutral distribution implicit in some banks' option prices, we have reached the conclusion that market participants fail to accurately price the likelihood of extreme events – *e.g.* a price crash on a bank's stock price. The estimation of the RN probability distribution allows us to assess the significance of this mispricing, to which we have tried to bring some explanations.

I.2/ Extracting the stock price risk neutral distribution from option prices

Let there be a public bank corporation which stock price at time T is equal to S_T. Assume that there is a market for European call options written on the bank's stock and consider a call option of strike k and maturity T.

As explained previously, assuming there are no arbitrage opportunities – or the milder condition *No Free Lunch with Vanishing Risk* (NFLVR), see Delbaen and Schachermayer [6] – the fair price of this European call is given by its expected discounted payoffs under the risk neutral measure:

$$c_k = e^{-rT}\mathbb{E}_{\mathbb{Q}}[(S_T - k)^+] = e^{-rT}\int_k^{\infty}(s-k)dF_{\mathbb{Q}}(s)$$

$$\Leftrightarrow\ c_k = e^{-rT}\left(\int_k^{\infty} s dF_{\mathbb{Q}}(s) - k\int_k^{\infty} dF_{\mathbb{Q}}(s)\right)$$

$$\Leftrightarrow\ c_k = e^{-rT}\left(\int_k^{\infty} s dF_{\mathbb{Q}}(s) - k\big[1 - F_{\mathbb{Q}}(k)\big]\right)$$

$\mathbb{Q}$ is the risk neutral measure or distribution and $F_{\mathbb{Q}}$ its cumulative distribution function. We are assuming that $f_{\mathbb{Q}}$ exists. Moreover, *r* is the discount rate, which we will assume to be constant.

Letting $f_{\mathbb{Q}}$ be the RN density, if we differentiate the above pricing integral c_k with respect to *k*, we obtain:

$$\frac{\partial c_k}{\partial k} = e^{-rT}\left(0 - kf_{\mathbb{Q}}(k) - 1 + F_{\mathbb{Q}}(k) + kf_{\mathbb{Q}}(k)\right)$$

$$\Leftrightarrow \frac{\partial c_k}{\partial k} = e^{-rT}\left(F_{\mathbb{Q}}(k) - 1\right).$$

Thus, we get an expression for the RN cumulative distribution function:

$$F_{\mathbb{Q}}(S_T) = 1 + e^{rT}\left.\frac{\partial c_k}{\partial k}\right|_{k=S_T} \qquad (1)$$

Note that, by definition: $\partial_k c_k \leq 0$.

By differentiating again with respect to strike, we obtain the RN density $f_{\mathbb{Q}}$:

$$f_{\mathbb{Q}}(S_T) = e^{rT}\left.\frac{\partial^2 c_k}{\partial k^2}\right|_{k=S_T} \qquad (2)$$

This expression, already noticed by Breeden and Litzenberger in 1978 [2], implies that the RN density associated to a price process S_T can be obtained by differentiating twice the pricing formula for a European Call, written on S_T, with respect to the strike *k*.

This result simplifies our statistical problem. We now have to estimate an option price positive function *g* of the strike *k* with the following form:

$$c_k = g(k).$$

The RN density will then be simply obtained by differentiating twice this function:

$$f_{\mathbb{Q}}(S_T) = e^{rT}\left.\frac{\partial^2}{\partial k^2}\right|_{k=S_T} g(k)$$

I.3/ Report objective and structure

The rest of this report deals with the statistical problem of estimating a RN distribution from option prices, assessing the accuracy of the estimation and interpreting the results obtained for a series of American banks, which have been designated as SIFIs.

The structure is as follows: in Section II, we describe the data we have used and the preliminary treatment we have applied to it. In Section III, we describe the technical methodology of RN distribution estimation, illustrated throughout the presentation with the example of a particular financial institution, Citigroup. We also discuss the results obtained for this bank by implementing a series of evaluation metrics. In Section IV, we compare results across four different financial institutions and expound some concluding remarks. In the appendix, we present some additional graphs, tables and results pertaining to Citigroup, as well as computations and results for other banks.

II – Data

II.1/ Data procedure

As explained above, our problem is to estimate the risk neutral density associated to the bank's stock price from its traded options' prices. We thus need real data for option prices.

It could be that, once we have at our disposal a sufficiently large sample of option prices, we could simply compute the RN density by calculating some numerical approximation of the 2nd derivative with respect to strike, for example using central or forward differences. However, two problems arise that make this approach unrealistic:

- First and foremost, there is simply not enough option prices. Strikes traded on the market are not separated by less than 1$, which is too large a step to do appropriate approximations of derivatives. We need to estimate a function which will enable us to generate further option prices, for strikes separated by 0.01$, 0.001$ or even 0.0001$.

- Secondly, as noted in [11] or [22], the implied volatility space is more conducive to estimation and interpolation techniques, while applying these same techniques to the option prices space could result in a less smooth and continuous RND. This phenomenon is explained by the greater dispersion of option prices relatively to implied volatilities: if we take the example of Citigroup, we observe in Tables 1 and 2 that the spread between the lower and the higher call mid-price retained for our analysis is 0.11-17.70, while this same spread for puts is 0.17-6.35; by contrast, as we will see later, implied volatilities are going to be comprised roughly in a 20%-50% interval.

To sum up, what we need is a sample of implied volatilities, to which we fit some function. This function will enable us to generate a vast range of implied volatilities. The last step will be to transform these implied volatilities into option prices that will then be used to estimate the RN density.

II.2/ Data description

The data we have used to establish the results of this report comes from OptionMetrics, a data company specialized in options market data from North America and the EU. We have been able to gather data for four different SIFI banks from the US: Citigroup, Bank of America Merrill Lynch (BAML), Goldman Sachs and Morgan Stanley.

For each bank, we have a series of option prices for American type options, both calls and puts, for the week going from April 7th 2014 to April 11th 2014 as well as for April 1st 2014. The prices given for each option contract are the best bid and the best offer at market closing on a particular day. The mid-price of the option is calculated as the average between the best bid and the best offer.

OptionMetrics also provides implied volatilities computed in-house. Their methodology is described more precisely in [16] but the main takeaway is that the implied volatility is determined for the option mid-price with a Cox-Ross-Rubinstein binomial tree and a continuous dividend yield.

OptionMetrics also provides traded volume information: for each option contract, we know the number of options that were traded across all exchanges.

The maturities of these options go from a few days to up to two years. In this report, results are computed using options maturing in around 3 months – on July 19th 2014.

As this point, some issues already arise. Indeed, the range of option mid-prices we can effectively use to estimate the RN density is relatively scarce: the RN distribution is an assessment about the market's expectation on the future price of the stock, thus we should only retain prices that we think are indicative of the situation of the market and agents' expectations.

It is widely acknowledged that the more liquid an asset is, the better its price reflects the current information about that asset. This observation implies that, in our estimation procedure, we should privilege data from the most traded options. To measure liquidity, given that a traded volume criterion would have drastically reduced the amount of exploitable data, we have devised a bid-offer spread criterion: retaining options for which this spread is too wide might yield an inaccurate assessment of the market's expectation, because it is generally assumed that the wider the spread is, the less liquid that particular option is.

The bid-offer spread criterion is also useful to evaluate the relevance of the option's mid-price. One could reasonably question whether a mid-price for a very wide bid-offer spread conveys any pertinent information about the market's expectation.

However, we also want our RN distribution to capture, at least partially, the tail behavior of the stock price, in order to have an idea of how the market evaluates the likelihood of extreme scenarios, such as a price crash. To do so, we need to include deep out-of-the-money (OTM) and deep in-the-money (ITM) options but, as we deviate further from the at-the-money (ATM) point, liquidity become scarcer and spreads wider. We have therefore to allow some flexibility in our choice of price options and consider a maximum bid-offer spread that is not too narrow.

We have to look at the bid-offer spreads of options and discard those for which this spread is too wide: to do so, we start by identifying the ATM strike; then both for ITM and OTM options we retain options until we reach the last option for which the bid-offer spread is narrower than 35% of the bid price, subject to the availability of an implied volatility figure. For illustrative purposes, in the two tables below we have inputted the data about Citigroup's options on April 7th 2014 that we have used to estimate its RN density, as well as traded volume data. Options that have been retained for estimating the RN density are highlighted. The closing price of Citigroup stock on that date was $46.55.

Although data selection has been undertaken under the basis of option prices, what we are really looking for are implied volatilities. Once we have selected options that respect our criteria, we are going to look at their implied volatilities as computed by OptionMetrics.

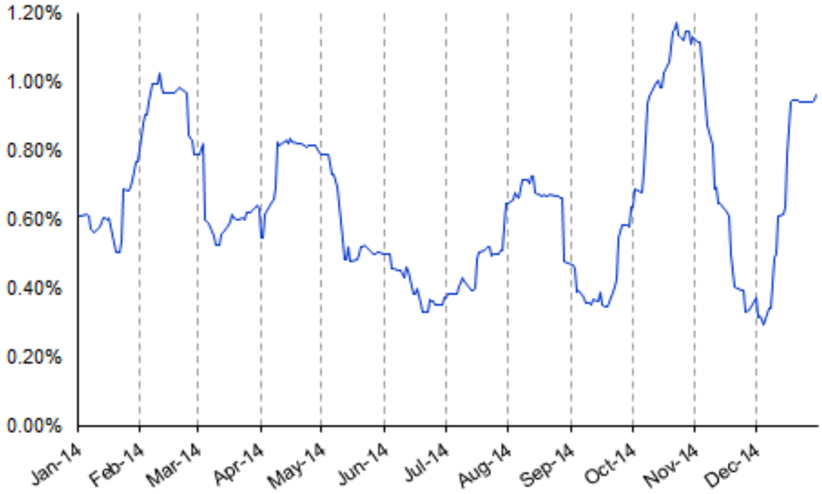
Calls				
Strike	*Bid-offer spread*	*Mid*	*Volume*	*Implied volatility*
25.0	16.5%	21.60	-	53.08%
26.0	18.9%	20.58	-	-
27.0	19.0%	19.55	-	-
28.0	19.8%	18.58	-	-
29.0	**13.3%**	**17.70**	**-**	**50.08%**
30.0	**13.4%**	**16.70**	**-**	**46.94%**
31.0	**14.3%**	**15.70**	**-**	**43.89%**
32.0	**13.0%**	**14.75**	**-**	**43.47%**
33.0	**15.2%**	**13.78**	**-**	**41.53%**
34.0	**17.4%**	**12.83**	**-**	**40.40%**
35.0	**17.1%**	**11.73**	**-**	**33.59%**
36.0	**20.4%**	**10.80**	**-**	**33.57%**
37.0	**20.8%**	**9.83**	**-**	**31.45%**
38.0	**12.3%**	**9.08**	**-**	**34.54%**
39.0	**10.5%**	**8.05**	**-**	**30.81%**
40.0	**2.9%**	**7.05**	**-**	**27.64%**
41.0	**7.6%**	**6.13**	**-**	**25.77%**
42.0	**9.9%**	**5.30**	**-**	**25.13%**
43.0	**3.3%**	**4.58**	**-**	**25.22%**
44.0	**2.6%**	**3.85**	**-**	**24.60%**
45.0	**3.2%**	**3.20**	**82**	**24.19%**
46.0	**1.5%**	**2.61**	**-**	**23.74%**
47.0	**1.4%**	**2.10**	**44**	**23.38%**
48.0	**1.8%**	**1.66**	**49**	**23.09%**
49.0	**2.3%**	**1.30**	**90**	**22.95%**
50.0	**3.1%**	**1.00**	**164**	**22.79%**
52.5	**4.3%**	**0.48**	**22**	**22.47%**
55.0	**4.8%**	**0.22**	**81**	**22.35%**
57.5	**10.0%**	**0.11**	**16**	**22.97%**
60.0	100.0%	0.08	1	25.08%
65.0	800.0%	0.05	-	29.52%

Puts				
Strike	*Bid-offer spread*	*Mid*	*Volume*	*Implied volatility*
25.0	-	0.03	-	-
26.0	-	0.04	-	48.49%
27.0	-	0.04	-	45.64%
28.0	-	0.04	-	42.88%
29.0	700.0%	0.05	-	41.63%
30.0	700.0%	0.05	-	38.97%
31.0	800.0%	0.05	-	36.95%
32.0	400.0%	0.06	-	35.37%
33.0	1000.0%	0.06	-	32.83%
34.0	500.0%	0.07	-	31.13%
35.0	175.0%	0.08	-	29.00%
36.0	114.3%	0.11	-	28.49%
37.0	**20.0%**	**0.17**	**4**	**28.24%**
38.0	**31.6%**	**0.22**	**1**	**27.37%**
39.0	**10.7%**	**0.30**	**1**	**26.59%**
40.0	**7.9%**	**0.40**	**1**	**25.87%**
41.0	**3.8%**	**0.53**	**16**	**25.27%**
42.0	**4.3%**	**0.71**	**19**	**24.71%**
43.0	**3.3%**	**0.93**	**34**	**24.16%**
44.0	**1.7%**	**1.20**	**16**	**23.64%**
45.0	**2.0%**	**1.55**	**73**	**23.25%**
46.0	**1.5%**	**1.96**	**23**	**22.84%**
47.0	**1.2%**	**2.45**	**92**	**22.54%**
48.0	**2.0%**	**3.02**	**8**	**22.40%**
49.0	**2.8%**	**3.65**	**38**	**22.11%**
50.0	**2.3%**	**4.35**	**6**	**21.88%**
52.5	**8.2%**	**6.35**	**-**	**21.43%**
55.0	12.0%	8.38	-	-
57.5	15.5%	10.78	-	-
60.0	23.3%	13.18	-	-
65.0	0.8%	18.43	-	-

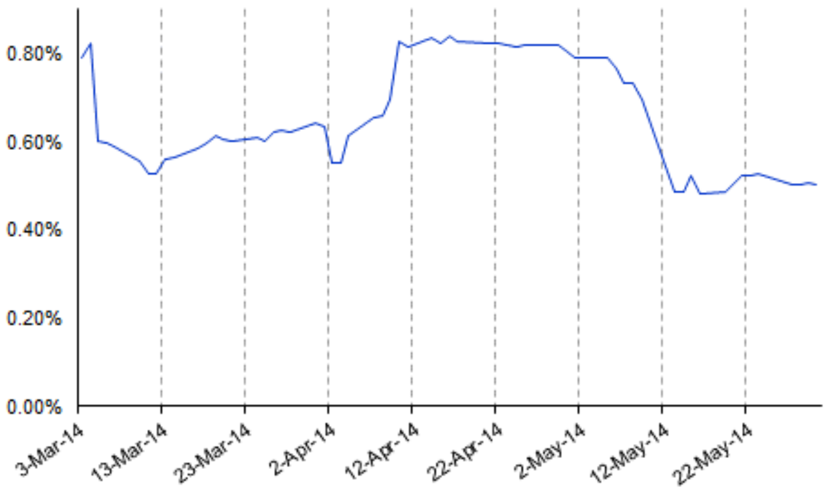
***Tables 1 and 2**: Option prices data for Citigroup on April 7th, for a 103 days maturity (2014).*
The selected prices are indicated in bold.

We notice here that we have retained for some strikes prices both call and put option <prices. This is done to avoid jumps in the implied volatility function: we have observed that the transition from calls' implied volatilities to puts implied volatilities might result in a jump, which is basically due to market imperfections but which nevertheless could distort our final results. Therefore, we have linearly interpolated OptionMetrics' implied volatilities for the range of strikes for which we have both call and put data.

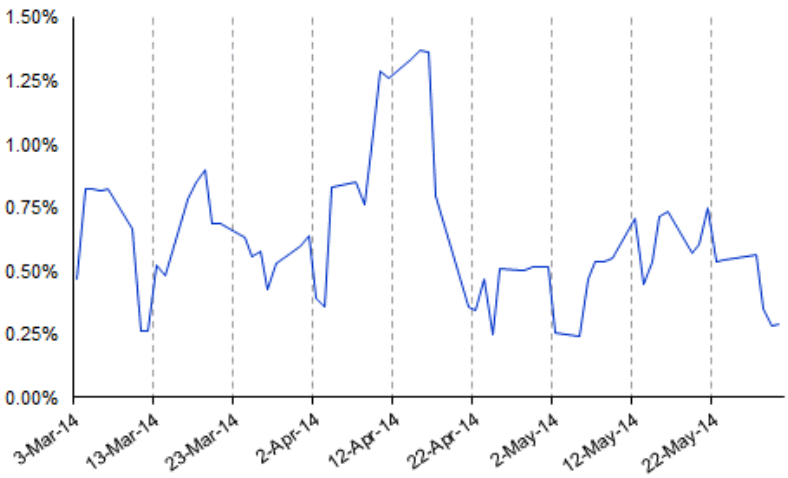
To put these prices in perspective, we have plotted a 21-day moving average of volatility of daily returns for the S&P 500 stock index, both for the full year 2014 and for the months of March, April and May, as well as a 5-day moving average only for March, April and May – on Graph 1, vertical lines indicate the month's beginning.



***Graph 1**: 21-day moving average of S&P 500 daily returns' volatility (2014)*



***Graph 2**: 21-day moving average of S&P 500 daily returns' volatility (March, April and May 2014)*



***Graph 3**: 5-day moving average of S&P 500 daily returns' volatility (March, April and May 2014)*

From our graphs, and recalling that our option prices are from April 7th, we conclude that there was a spike of volatility at the time our prices were quoted. There were 4 other similar volatility spikes in 2014 – in January/February, in July, in September/October and in December.

Nevertheless, this spike remains moderate: if we look at Graph 1 and Graph 2, we observe the 21-day moving average was around 0.65% on April 7th; however this statistic fluctuated between 0.30% and 1.17% in 2014 hence the April 7th figure was on the middle of the range. Moreover, the spikes that occurred in January/February, September/October and December seem much sharper than our April spike. Finally, the 5-day average volatility was 0.85% on that day, against a minimum and a maximum of 0.06% and 1.82% respectively on 2014, so again the April 7th value is on the middle of the range.

From the above graphs we conclude that our option price data is extracted from a period not subject to any particular market stress. These option prices should hence not incorporate any premium related to abnormal market volatility and their pricing should be rational.

II.3/ Converting implied volatilities into option prices

Once the first phase of the data procedure – obtaining implied volatilities – is completed, we need to fit a curve to these volatilities, generate new volatility data and then convert it back into option prices. We describe in detail the function fitting methodology in the next section. Here we briefly explain how we transform implied volatilities into option prices and present the additional data we need to do this.

In order to implement this conversion, we use the Black-Scholes framework [1] – alternatively, another option pricing model can be used, such as the Heston model [18]. Indeed, this is necessary if we want to have at our disposal a computational tool to convert option prices into implied volatilities and the other way round. This trick does not imply that we are assuming the Black-Scholes framework: it is a mere tool used to ease computations and improve their accuracy – we have explained previously the advantages of going through the implied volatility space.

To implement Black-Scholes, we need additional data. First, we need to input into the formula a risk free rate. We have chosen the quarterly US Dollar Libor, which traded on average at 0.227% during April 2014 (*source: global-rates.com*).

Secondly, we also need a dividend yield to evaluate European option prices. The best choice would be to extract an implied dividend yield from market prices by exploiting call-put parity for European options, as we would obtain market consistent prices for European options. Letting $c(k,s_0,T)$ be the price of an European call with strike k and maturity T written on a stock which price today is s_0 - $p(k,s_0,T)$ being the price for a put – f_0 the price of a forward contract on this same stock with the same maturity and d the dividend yield, we have the following equalities:

$$c(k, s_0, T) - p(k, s_0, T) = (f_0 - k)e^{-rT}$$

$$f_0 = s_0 e^{(r-d)T}.$$

From these two equations we can derive a formula for the implied dividend yield:

$$d_I = r - \frac{1}{T}\ln\left\{\frac{[c(k, s_0, T) - p(k, s_0, T)]e^{rT} + k}{s_0}\right\}.$$

However, as we have pointed out, we do not have European option data, thus we cannot extract implied dividend yields.

Instead, in this paper we have calculated dividend yields using historical data. Two approaches are possible:

- *Annual dividend yield*: dividends paid on year *N* are assumed to be linked to a company's income from year *N-1*. As income for year *N-1* is known at the beginning of year *N*, markets participants anticipate dividend payments during year *N* based on that figure and the stock price depreciates throughout year *N* at the dividend yield rate *d*. In this paper we have retained this interpretation.
  To compute *d*, we have recorded dividends paid by banks in 2014 – $\mathrm{Div}_{1,N}$, $\mathrm{Div}_{2,N}$, etc. – and divided their sum by the average stock price in 2013 – $\bar{s}_{2013}$ – computed simply as the average between 2013's opening price and 2013's closing price.

$$d = \frac{\sum_i \mathrm{Div}_{i,N}}{\bar{s}_{N-1}}$$

- *Option lifetime dividend yield*: the dividend yield *d* is based on dividend payments during the lifetime of the option – $\mathrm{Div}_{1,[t,T]}$, $\mathrm{Div}_{2,[t,T]}$, etc. – such that the depreciation of the stock price arising from these payments is distributed along the option's lifetime – between the current date *t* and the expiration date of the option *T*. The rate can be annualized by dividing by *T-t*.

$$d = \frac{1}{T-t} \times \frac{\sum_i \mathrm{Div}_{i,[t,T]}}{s_{t_0}}$$

Data can be found in Tables 3 and 4.

Bank	Citigroup	BAML	Goldman Sachs	Morgan Stanley
Year 2013 opening price	\$40.91	\$12.05	\$131.30	\$20.16
Year 2013 closing price	\$52.11	\$15.57	\$177.26	\$31.36
2013 average yearly price	\$46.51	\$13.81	\$154.28	\$25.76
April 1st, 2014 (closing)	\$47.80	\$17.34	\$165.92	\$31.21
April 7th, 2014 (closing)	\$46.55	\$16.38	\$158.56	\$29.52
April 8th, 2014 (closing)	\$46.60	\$16.44	\$156.56	\$29.53
April 9th, 2014 (closing)	\$47.16	\$16.62	\$158.16	\$30.22
April 10th, 2014 (closing)	\$46.23	\$16.12	\$155.98	\$29.25
April 11th, 2014 (closing)	\$45.68	\$15.77	\$152.72	\$28.47

Table 3*: Relevant stock prices for the four banking corporations (2013, 2014)*

Bank	Citigroup	BAML	Goldman Sachs	Morgan Stanley
2014 dividends	4 × \$0.01	2 × \$0.01 2 × \$0.05	3 × \$0.55 1 × \$0.60	1 × \$0.05 3 × \$0.10
Dividend paid during option lifetime	1 × \$0.01 (May 1st)	1 × \$0.01 (June 20th)	1 × \$0.55 (May 28th)	1 × \$0.10 (April 28th)
Annual dividend yield	0.086%	0.869%	1.458%	1.359%
Annualized option lifetime dividend yield	0.076%	0.193%	1.110%	1.073%

Table 4*: Dividends and dividend yields for the four banking corporations (2014)*

We observe that the difference between annual dividend yields and annualized option lifetime dividend yields is negligible, particularly considering the general weakness of these dividend rates. Choosing one or the other interpretation should not have any meaningful impact on our results. Furthermore, as pointed out in [16], the sensitivity of the theoretical option price to the dividend yield is small except in the case of long-dated options. Thus we conclude that our approach to the dividend yield is reasonable.

Once we have gathered all the required data, we can convert implied volatilities into option prices. We use in R the `RQuantLib` package to compute them.

III – Risk Neutral Distribution of stock prices: methodology and results

Once we have rigorously delimited and constructed the data which we will use to estimate the RN distribution, we can begin to apply statistical interpolation and smoothing techniques to extract the RN density.

The first step, as explained, is to construct a function that relates option strikes to implied volatilities. Given a series of strikes (k_i), we are going to look at two different kinds of function:

- A deterministic modelling of the implied volatility function :

$$\hat{\sigma} : k \rightarrow \hat{\sigma}(k)$$

- A probabilistic modeling of implied volatility, by considering it as a random variable and modelling its expectation conditional on the strike K:

$$\widehat{\mathbb{E}}[\sigma \mid K] : k \rightarrow \widehat{\mathbb{E}}[\sigma \mid K = k]$$

Once we have determined a suitable function, we need to convert implied volatilities back into option prices. Although our initial options were American, once we have estimated implied volatilities we can, assuming that for example the Black-Scholes model holds, obtain European option prices by inputting our parameters – dividend yield, risk free rate, etc. – and the estimated implied volatilities into the model. Indeed, this is necessary as the formula that allows us to get the RN density only holds for a contingent security with a single payoff date – therefore we need European option prices and not American option prices.

In proceeding as described in the previous paragraph, we are assuming that American and European options' implied volatilities are equal. This assumption seems reasonable to us, as the implied volatility is a parameter than should pertain to the stock price and not the option price.

We are therefore implicitly assuming that the difference between the price of an American call and a European call is fully explained by the peculiarity of the American option, which allows the holder to exercise it at any moment in time between the current date and the maturity, instead of being explained by differences in implied volatilities.

The following step consists of estimating the RN density: given the option prices data we have determined on the previous section, if we have enough prices, we can then approximate the RN density using 2nd order central differences of option prices. At this point, we might have to deal with some irregularities arising during the estimation procedure.

Finally, we compute the resulting RN density by averaging all the estimated RN densities. We then fit it to a known parametric probability distribution, and test whether this approximation is valid.

To illustrate the whole procedure, we show various results for Citigroup throughout the text.

III.1/ Estimating the implied volatility function

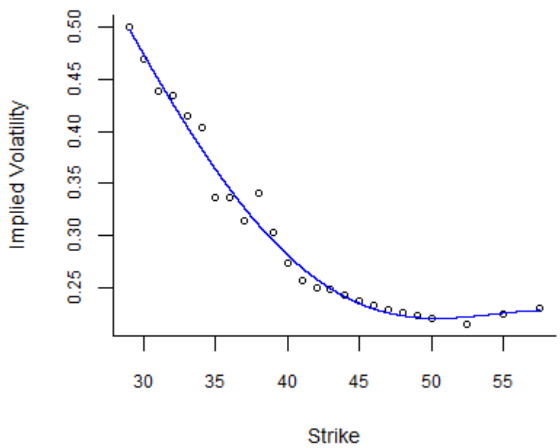
III.1.A/ Polynomial Interpolation

The first method is polynomial interpolation. In [11], Figlewski argues that this highly tractable technique yields very satisfactory results. Its main advantage lies in its simplicity, as the implied volatility function is interpolated by a single polynomial. We choose a 4th degree polynomial, such that the implied volatility function is:

$$\hat{\sigma}(k) \; = \; \beta_0 + \beta_1 k + \beta_2 k^2 + \beta_3 k^3 + \beta_4 k^4.$$

For this technique, choosing 4th degree seems optimal. Indeed, as the implied volatility curve should not fluctuate, adding further degrees could result in overfitting, especially given that our dataset is limited in size.

In Graph 4, we show the results of the polynomial fit. The circles represent real data.

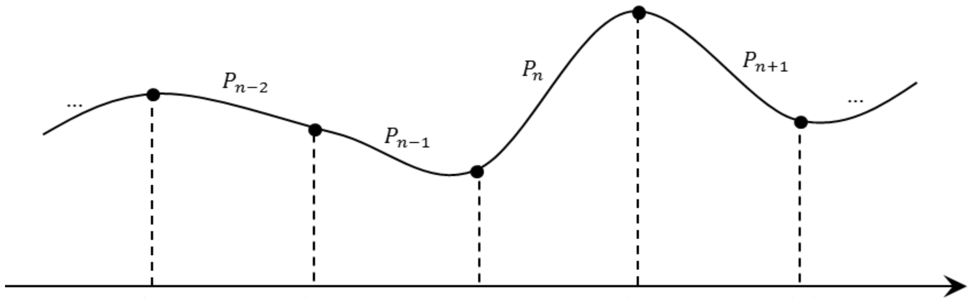


Graph 4*: Polynomial Interpolation of Implied Volatility (Citigroup, April 7th 2014)*

The resulting interpolated curve seems to capture accurately the implied volatility structure, especially as we get closer to the ATM region. Square errors analysis is undertaken in page 17.

III.1.B/ Smoothing Splines

The second technique consists in a sophistication of polynomial interpolation: smoothing polynomial splines. In such a model, the implied volatility function is modelled by a set of polynomials of the same degree, each one describing the curve on a particular interval delimited by knots:

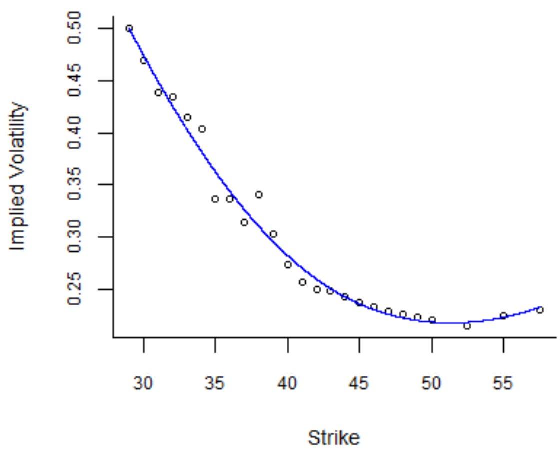


There exists a rich literature on splines and various techniques are available. In this report, we have chosen smoothing splines, the same technique used by Lai in [22]. Smoothing splines are determined by resolving the following optimization function:

$$\min_{\hat{\sigma}}\left[\sum_{i=1}^{n}\left(\sigma_i-\hat{\sigma}(k_i)\right)^2+\alpha\int_{-\infty}^{+\infty}\hat{\sigma}''(k)^2dk\right]$$

The second term of the expression ensures that the function is sufficiently smooth, and the α parameter enables the user to determine the importance given to smoothing. In such a spline technique, knots are automatically selected to optimize smoothing – there exist other methodologies, such as interpolation splines, in which the user can chose where to place the knots; such an approach is used for instance by Figlewski in [11].

As for polynomial interpolation, we choose 5th degree polynomials to fit the implied volatility function. To do so, in our implementation we use the R package `pspline`, which allows to choose the degree of the smoothing spline. This package also allows the user to choose an automatic selection criterion for the smoothing parameter alpha; in this case, we have retained the cross-validation criterion. The fitted curve can be found in Graph 5.



Graph 5: *Spline Smoothing of Implied Volatility (Citigroup, April 7th 2014)*

III.1.C/ Nadaraya-Watson Estimator

The third technique used to estimate the RN density is kernel regression, through the Nadaraya-Watson estimator. In such a setting, the target function is the expectation of implied volatility conditional to the strike.

From a probability viewpoint, the conditional expectation of a random variable *Y,* given that the value taken by a random variable *X* is *x*, is equal to:

$$\mathbb{E}[Y|X=x] = \frac{1}{f_X(x)} \int y f_{X,Y}(x,y)dy.$$

The function *f* refers either to the probability density of *X* or to that of the couple (*X, Y*).

By estimating densities through a Parzen-Rozenblatt kernel *f*, an estimator of the expectation of implied volatility conditional to the strike is the Nadaraya-Watson estimator:

$$\widehat{\mathbb{E}}[\sigma \mid k] = \frac{\sum_{i=1}^{n}\left(f\left(\frac{k-k_i}{h}\right)\sigma_i\right)}{\sum_{i=1}^{n}\left(f\left(\frac{k-k_i}{h}\right)\right)}$$

In this report, we take the Gaussian kernel, which we deem sufficiently smooth for our purpose – other kernels collapse to 0 when $k - k_i > h$, which might cause trouble given the sparsity of our data:

$$f(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}$$

The estimator can thus be rewritten as follows:

$$\widehat{\mathbb{E}}[\sigma \mid k] = \sum_{i=1}^{n}\left(\frac{e^{-\frac{1}{2}\left(\frac{k-k_i}{h}\right)^2}}{\sum_{j=1}^{n} e^{-\frac{1}{2}\left(\frac{k-k_j}{h}\right)^2}}\right)\sigma_i = \sum_{i=1}^{n} \sigma_i w_i(k).$$

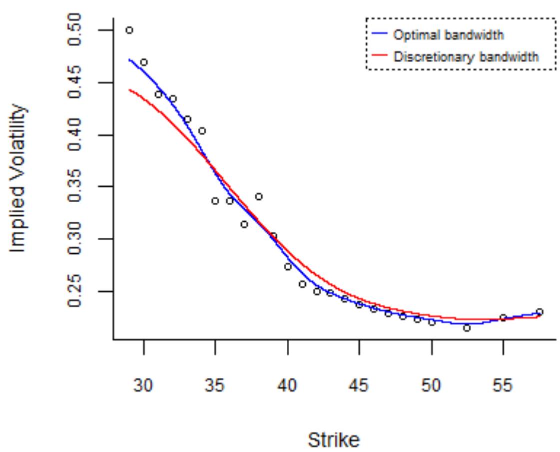
The estimated implied volatility is equal to a weighted sum – alternatively, it can also be rewritten as a weighted average – of the observed implied volatilities, where the weights depend on the information – strike *k*.

The main difficulty in Nadaraya-Watson regression is the optimal selection of the bandwidth parameter *h*, that arbitrages between the estimation bias and the variance. Given that we are using the Gaussian kernel, we have retained Silverman's rule-of-thumb [28] to select *h*:

$$h^* \approx 1.06 \times SD(k) \times n^{-1/5} \approx 4.45$$

However, we observed a posteriori some undesirable properties on the estimated RN density curve, therefore we decided to make a second estimation of the implied volatility function by setting the bandwidth *h* equal to 9, which will yield a smoother RN density.

For implementation in R, we have used the `ksmooth` function. The fitted curve is hereafter.



***Graph 6**: Comparison of Nadaraya-Watson estimators for Implied Volatility (Citigroup, April 7th 2014)*

As we consider that the RN density which results from using the Nadaraya-Watson optimal bandwidth estimator is not smooth enough – there were too many bumps in the curve – we have discarded this result and retained only the estimated function obtained by fixing the bandwidth equal to 9.

III.1.D/ Comparing results

After having estimated all four implied volatility functions, we have computed the sum of squared errors between our initial implied volatility data and the estimated implied volatility – only for strikes for which we have empirical data. Results can be found in Table 5.

Method	Polynomial interpolation	Spline smoothing	Nadaraya-Watson estimator: h = h*	Nadaraya-Watson estimator: h = 9
SSE for implied volatility estimation	0.0031028	0.0031403	0.0032261	0.0093162

***Table 5**: Sum of Squared Errors for all implied volatility estimation methods*

The first observation is that Nadaraya-Watson estimators perform worse than polynomial-based estimators. However, recall that at the beginning of the section we stated that we were going to look at two different types of functions, one relating to the implied volatility value:

$$\hat{\sigma} : k \rightarrow \hat{\sigma}(k)$$

the other one relating to the implied volatility conditional expectation:

$$\widehat{\mathbb{E}}[\sigma \mid K] : k \rightarrow \widehat{\mathbb{E}}[\sigma \mid K = k]$$

The two polynomial-based estimators correspond to the first class, while the Nadaraya-Watson estimator looks at conditional expectations. In that sense, it seems natural that the first two methods are more accurate than those derived from the Nadaraya-Watson procedure as the latter tries to model the average of a function, whereas the former represents in the function itself.

Overall, the best performer is the polynomial interpolation, although the difference with respect to spline smoothing is negligible. The Nadaraya-Watson estimator with optimal bandwidth is also close to the polynomial-based methods, however as we stated above this is achieved at the cost of RN density's smoothness. Finally, the Nadaraya-Watson estimator with discretionary bandwidth yields a SSE which is more than triple that the SSE of the polynomial interpolation.

III.2/ From implied volatilities to option prices

Once we have an estimation of implied volatilities, we need to get back to the option prices space.

As it has already been explained, in this report we have retained the Black-Scholes option pricing model for conversion purposes. In this framework, the price of a European call option is given by:

$$c_k = e^{-rT}\mathbb{E}_{\mathbb{Q}}[(S_T - k)^+] = S_0 e^{-Td}\varphi(d_1) - ke^{-Tr}\varphi(d_2)$$

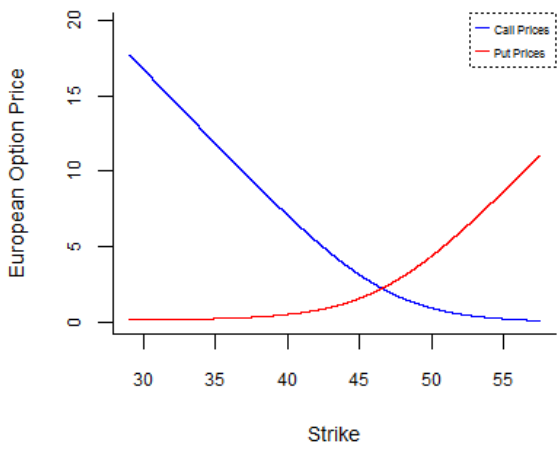
$$d_1 = \frac{1}{\sigma\sqrt{T}}\left[\ln\left(\frac{S_0}{k}\right) + \left(r - d + \frac{1}{2}\sigma^2\right)T\right]$$

$$d_2 = d_1 - \sigma\sqrt{T}$$

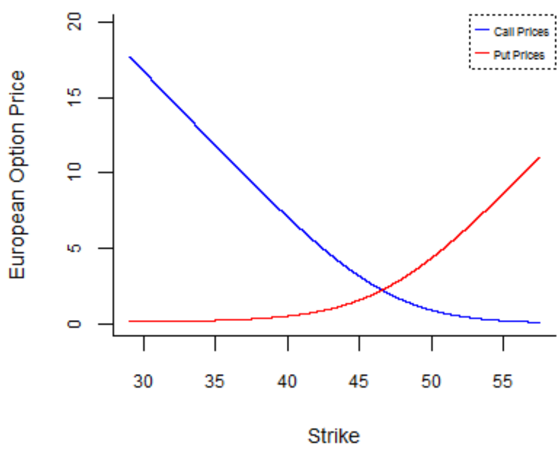
where φ denotes the standard normal cumulative distribution.

As pointed out by Figlewski [11] and Lai [22], the use of the Black-Scholes model is simply a computational tool to shift between the implied volatility space and the option prices space. One immediate observation would be that, by using the Black-Scholes function as a transforming device between volatilities and prices, we are thereby setting the risk neutral measure to be a log-normal distribution. However, this is not accurate: indeed, the Black-Scholes model assumes that volatility is constant, whereas our project by definition assumes that this is not the case – as it can be checked instantly by looking at the above graphics. Therefore, as highlighted by Figlewski, we do not get automatically a log-normal distribution when transforming back implied volatilities into option prices.

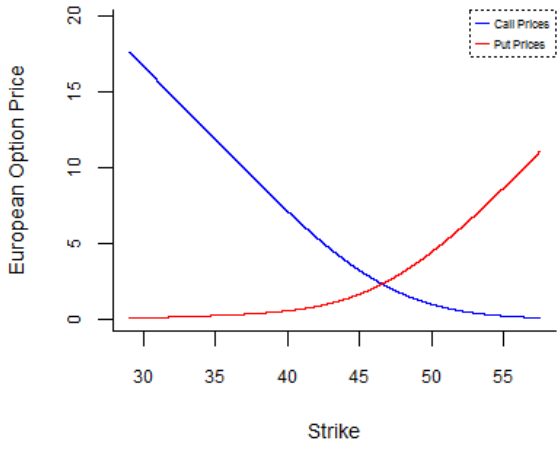
We use again the `RQuantLib` package to transform our estimated implied volatilities into option prices. We have done this both assuming our options are calls and puts. The results can be found in Graphs 7, 8 and 9, where the blue curve is evidently the call price curve, while the red one is the put curve. As we can observe, there is no evident difference between each one of the three methods.



***Graph 7**: European option prices computed from implied volatilities obtained by polynomial interpolation (Citigroup, April 7th 2014)*



***Graph 8**: European option prices computed from implied volatilities obtained by spline smoothing (Citigroup, April 7th 2014)*



Graph 9*: European option prices computed from implied volatilities obtained by the Nadaraya-Watson estimator h = 9 (Citigroup, April 7th 2014)*

III.3/ Estimating the risk neutral probability

III.3.A/ Estimation procedure and results

Now that we have our set of estimated option prices, we can finally extract both the RN cumulative distribution function and the RN density. To do so, we proceed by discretization in (1) as for instance in Fisglewski [11] with a 2-step approach. The cumulative distribution can be approached using 1st order central differences: as the cumulative distribution function is the 1st derivative of option price with respect to the strike, it can be approximated for a particular strike k_n, given call option prices (k_i), through the following formula:

$$F_{\mathbb{Q}}(k_n) \approx e^{rT}\left(\frac{c_{n+1}-c_{n-1}}{2\Delta k}\right)+1.$$

where Δk denotes the difference between two consecutive strikes.

In the case of put option prices (p_i):

$$F_{\mathbb{Q}}(k_n) \approx e^{rT}\left(\frac{p_{n+1}-p_{n-1}}{2\Delta k}\right).$$

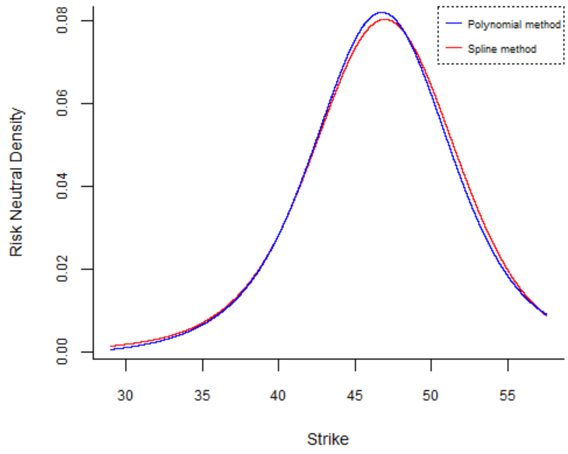
For approaching the RN density, we use 2nd order central differences; the formula is the same for both calls and puts, so given a set of option prices (o_i):

$$f_{\mathbb{Q}}(k_n) \approx e^{rT}\left(\frac{o_{n+1}-2o_n+o_{n-1}}{(\Delta k)^2}\right).$$

We set the difference between consecutive strikes Δk equal to 0.0001. In addition, in the case of Citigroup – on April 7th 2014 – after estimating the implied volatility function, we compute 285,001 different implied volatilities for strikes separated by $0.0001. Using the Black and Scholes formula, we obtain 285,001 option prices, which constitutes a sufficiently dense set of values to obtain satisfactory results.

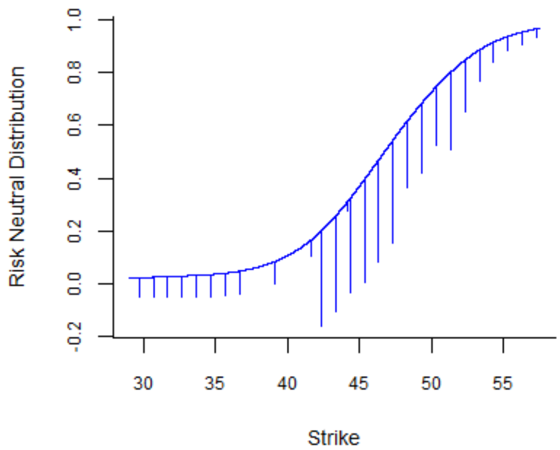
Figures for the cumulative distribution and the density function for polynomial interpolation and spline smoothing can be found in the appendix. In Graph 10 we show in the same plot both density functions.

As the difference between these functions computed with call prices and put prices is negligible – which is reassuring, otherwise we would have different probability measures for calls and puts, a phenomenon that would not make much sense – we only display curves computed with put option prices – we will explain later why we decided to keep the density estimated with put prices.

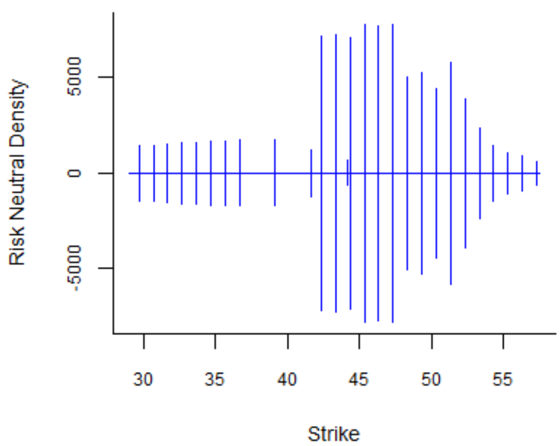


Graph 10*: Comparison of RN densities for the polynomial interpolation and the spline smoothing methods (Citigroup, April 7th 2014)*

When estimating the distribution and the density functions with the option prices obtained through the Nadaraya-Watson estimator we have encountered some technical issues. The first curves we get by implementing central differences methods are the following:



Graph 11*: Risk neutral distribution for Nadaraya-Watson estimator h = 9 (Citigroup, April 7th 2014)*



Graph 12*: Risk Neutral Density for Nadaraya-Watson estimator h = 9 (Citigroup, April 7th 2014)*

It seems that there is some periodic numerical instability related to our estimated data. We also encountered it when estimating the functions for the other banks.

To solve it, we have redefined each point of the cumulative function and the density function as follows:

$$x_i = x_i \mathbb{1}_{(x_i - x_{i-1} < \varepsilon)} + x_{i-1} \mathbb{1}_{(x_i - x_{i-1} > \varepsilon)}.$$

When the difference between a point estimate and its predecessor is greater than ε, we change the value of point x_i by the value of the previous point x_{i-1}. To determine the value of ε, we first compute difference between our points:

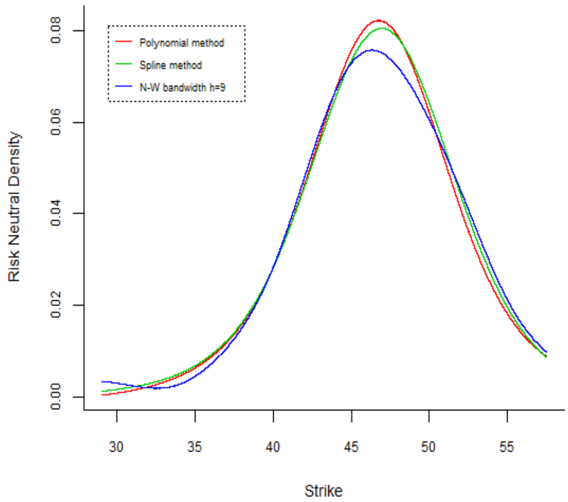
$$\Delta x_i = x_i - x_{i-1}$$

We then sort these differences by ascending order and we compute jumps:

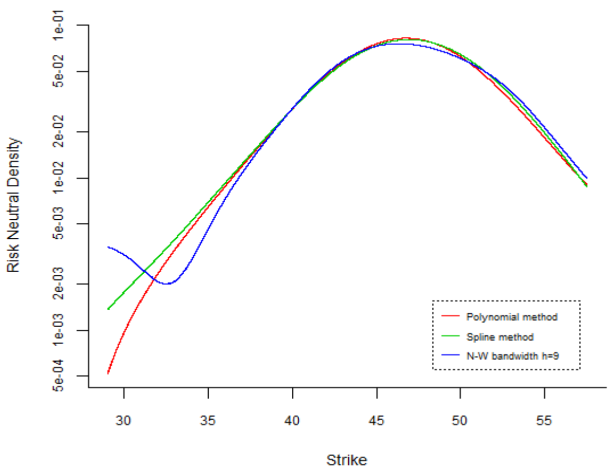
$$J_i = \Delta x_{(i)} - \Delta x_{(i-1)}.$$

ε is then simply equal to the maximal jump: $\varepsilon = \max_i J_i$.

In our case we have chosen ε equal to 0.028 for the cumulative function and to 1 for the density. This method allows us to extract all aberrant values. Amended cumulative distribution and density functions can be found in Appendix A. We have finally plotted all three RN densities in a single graph for better comparison.



***Graph 13**: Comparison of RN densities (Citigroup, April 7th 2014)*



***Graph 14**: Comparison of RN densities, logarithmic scale (Citigroup, April 7th 2014)*

In Graph 14 we have plotted these same distributions, but using a logarithmic scale to emphasize the tail behavior.

The reader can notice that the Nadaraya-Watson function does not suffer from numerical instability any longer, however by taking a logarithmic scale we observe a sharp variation on the left tail – we will discuss this below.

Both the polynomial and the spline based methods seem to deliver similar results, which is consistent with the fact that both techniques use polynomials to approximate the function. As pointed out in [11], relatively simple techniques can yield results very close to those obtained with more sophisticated methods, which seems to be confirmed by the similarity between these two curves.

III.3.B/ Results assessment

In order to explore the accuracy of our estimations, we undertake the following test. Given that we are producing estimates of probability densities, if the estimate captured the whole behavior of stock price dynamics then the value of the integral of the density should be equal to 1:

$$I_{f_{\mathbb{Q}}} = \int_{-\infty}^{+\infty} f_{\mathbb{Q}}(s)ds = 1$$

To estimate the integral's value, we use a rectangular method and construct the following estimator:

$$\hat{I}_{f_{\mathbb{Q}}} = \sum_{i=1}^{N} \hat{f}_{\mathbb{Q}}(k_i) \times \Delta k = \Delta k \sum_{i=1}^{N} \hat{f}_{\mathbb{Q}}(k_i)$$

$$\forall i \in [1; N-1]: \Delta k = k_{i+1} - k_i.$$

$\hat{f}_{\mathbb{Q}}$ is our estimator of the RN density – either through polynomials, splines or the Nadaraya-Watson method – k_i are the strike series we generated to evaluate the density, *Δk* is equal to \$0.0001 and *N* to 285,001 – see above.

We would like to have at least an estimated value of this integral as close to 1 as possible, while being inferior to 1. Indeed, a way of interpreting results is to consider the obtained percentage –probability – as an indication of the likely risk neutral future behavior of the stock price. We are using strikes from \$29 to \$57.5, while the stock price is \$46.55, which implies respectively a -37.7% and +23.5% return over a 3-month period. It seems reasonable that the integral value between 29 and 57.5 should be rather close to one, while remaining lower than one to allow for more extreme events to happen.

Results are showed in Table 6.

Method	Polynomial interpolation	Spline smoothing	Nadaraya-Watson estimator: h = 9
Density integral estimate	95.42%	96.62%	94.64%

***Table 6**: Density integral estimates for polynomial, spline and Nadaraya-Watson methods*

Overall, results are satisfactory as all three integral estimates are very close to 1 while leaving around a 5% RN probability to extreme events. The Nadaraya-Watson method is the one that assigns the higher likelihood to tail events, which is expected seing the strange tail behavior it manifests. The difference between the three methods is nonetheless negligible.

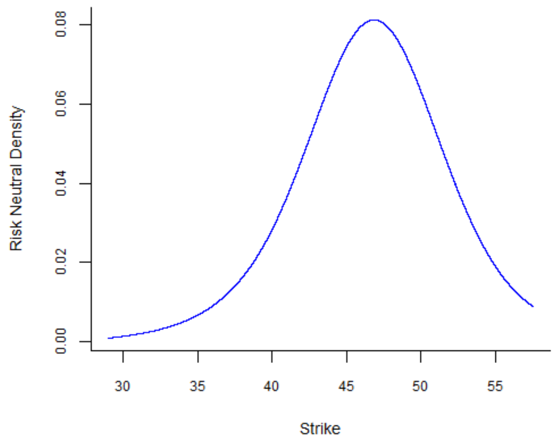
III.4/ Fitting a known parametric probability distribution

III.4.A/ Fitting procedure

Now that we have multiple estimates of the true RN density of the stock price process, it is interesting to fit a parametric distribution to these estimates. Such an approach allows us to derive closed-form formulas for the VaR for instance.

In Graph 14 above, we observe that the Nadaraya-Watson density has an undesirable feature in its left tail. This bump could distort results by giving that tail much more weight than plausible.

Therefore, we have computed average risk neutral distribution and density using only the densities estimated by polynomial interpolation and spline smoothing. The resulting RN density is shown on Figure 15.



***Graph 15**: Average RN density (Citigroup, April 7th 2014)*

It is worth highlighting again the negligible difference that we find between the distributions estimated with call or put prices. For example, for the average density, the sum of square errors between the average density computed with call prices and the one obtained through put prices is of the order of $O(10^{-8})$.

Then we have tried to fit a non-standardized Student distribution to the average estimated RN density. The advantage of the non-standardized Student distribution is that it has some of the nice characteristics of the Gaussian distribution, while allowing for fatter tails, a feature which is consistent with empirical observations on stock prices movements. The mentioned distribution is characterized by three parameters: its degrees of freedom η, a location parameter μ and a scale parameter σ. Let t_S be a random variable distributed according to a Student distribution with η degrees of freedom, then the random variable X below follows a non-standardized Student distribution with η degrees of freedom and with location and scale parameters μ and σ respectively:

$$X = \mu + \sigma t_S$$

The density function of the non-standardized Student distribution is given by the following formula, where Γ stands for the Gamma function:

$$f(x) = \frac{\Gamma\left(\frac{\eta+1}{2}\right)}{\Gamma\left(\frac{\eta}{2}\right)\sqrt{\pi\eta\sigma^2}}\left(1+\frac{1}{\eta}\left(\frac{x-\mu}{\sigma}\right)^2\right)^{-\frac{\eta+1}{2}}.$$

Given that we do not have a sample realization of a random variable but an estimated density function, instead of implementing a probabilistic technique such as maximum likelihood we have simply minimized the sum of squared errors between a set of parametrized non-standardized Student distributions and our averaged estimate of the RND. We set a constraint on the location parameter: we know that the expectation of the Student-distributed variable t_S is 0, thus:

$$\mathbb{E}[X] = \mu + \sigma\mathbb{E}[t_S] = \mu.$$

Moreover, under the risk neutral distribution, the expected return of a stock S_T is equal to the risk-free rate diminished of the dividend yield *d*:

$$\mathbb{E}_{\mathbb{Q}}[S_T] = s_0 e^{T(r-d)}$$

Therefore we set the following constraint:

$$\mu = s_0 e^{T(r-d)}.$$

As stock prices cannot take negative values, there is a limitation of choosing a model based on a distribution with support on $\mathbb{R}$. Nevertheless, we can reasonably assume that any company has a non-zero probability of default, which could be read as hypothetical negative prices for the shares. That is why we introduce a non-standardized Student distribution to model the stock prices, such that the probability of default corresponds to the probability that the random variable associated with this Student distribution takes negative values.

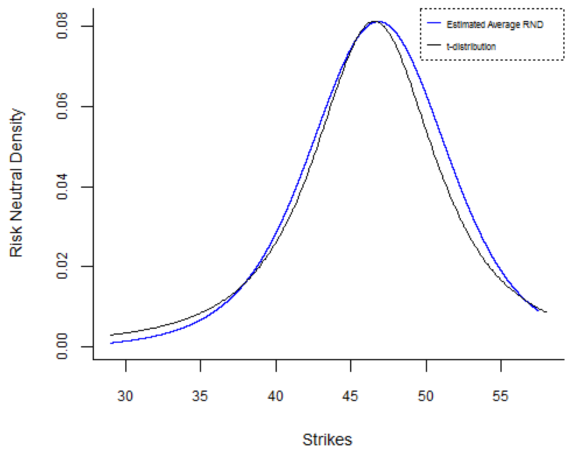
In Citigroup's case and given that we do not have a continuous rate, we obtain:

$$\mu = 46.55 \times \big(1 + 0.282 \times (0.227\% - 0.086\%)\big)$$

$$\Leftrightarrow \mu = 46.55 \times (1 + 0.282 \times 0.141\%) \approx \$46.5685.$$

Our fitting procedure consists then of trying different values for scale parameters, setting the degrees of freedom equal to 2, 3, 4, 5, 6, 7 and 8 in order to have relatively fat tails.

First, we have visually identified Student distributions that were close to our estimate. For example, in case η = 2, the value for the parameter is σ = *4.35*.



***Graph 16**: First approximation by a non-standardized Student distribution with 2 degrees of freedom (Citigroup, April 7th 2014)*

Then:

- We compute the sum of logs squared errors for 2501 non-standardized Student distributions, varying the scale parameter from 66.66% to 133.33% of the visual approximation value, *i.e.* σ = *4.35* above. Our objective function is the following:

$$\sum\left(\log(\hat{x}_i) - \log\left(x_i^{ts}\right)\right)^2.$$

Using logarithms allows us to give more relative weight to the tails of the distribution during the fitting procedure: given that the values of the estimated density – $\hat{x}_i$ – and the Student density – x_i^{tS} – are below one, taking the logarithm increases sharply the value of the density at the tails. For example, if x = *0.1001* and y = *0.1000*, the difference between x and y is 0.0001 and the difference between their logarithms is approximately 0.0010; however if x = *0.0011* and y = *0.0010*, their difference is still equal to 0.0001 but the difference between their logarithms increases to 0.0953 – almost ten-fold;

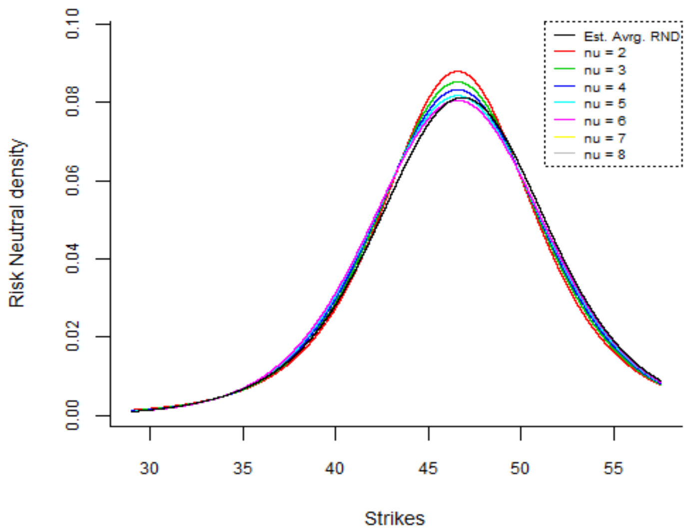
- We pick the best fit from above;
- We compute the sum of logs squared errors for 2501 non-standardized Student distributions, varying the scale parameter from 6.66% to 103.33% of the above approximation value.

We found the following optimized parameters for Citigroup on April 7th, 2014:

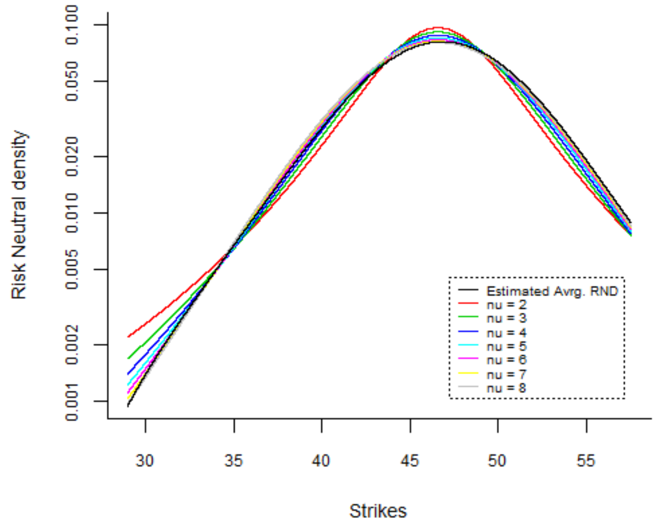
Degrees of freedom	***σ*** *(visual approximation)*	***σ*** *(optimized fit)*
2	\$4.35	\$3.6681
3	\$3.90	\$4.0148
4	\$4.35	\$4.2669
5	\$4.50	\$4.4547
6	\$4.65	\$4.5990
7	\$4.70	\$4.7134
8	\$4.75	\$4.8058

Table 7: *Optimal scale parameters for fitted non-standardized Student distributions*

Graphs 17 & 18 represent the RN density estimate as well as the fitted Student distributions. On Graph 18, the scale is logarithmic.

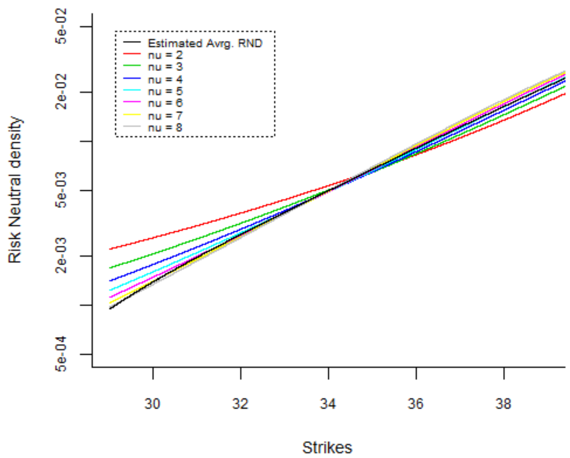


Graph 17*: Comparison of average RN density and fitted non-standardized Student distribution (Citigroup, April 7th 2014)*

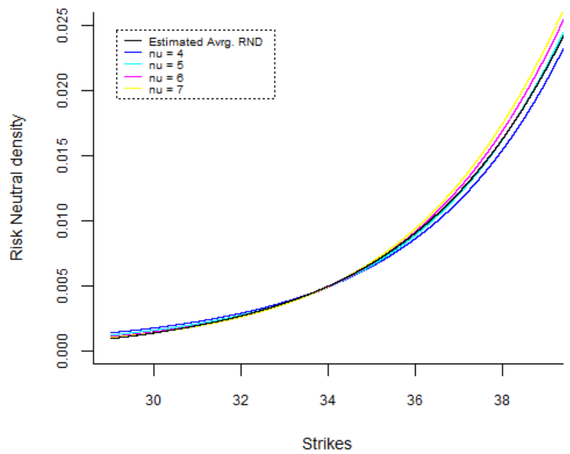


***Graph 18**: Comparison of average RN density and fitted non-standardized Student distributions, logarithmic scale (Citigroup, April 7th 2014)*

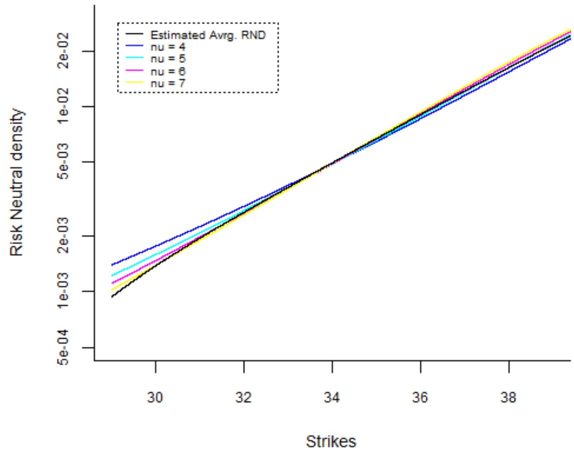
We now zoom in the left tail of the distribution, showing only strikes between \$29 and \$39. We first show all distributions on a logarithmic scale; then on Graphs 20 and 21 we only show distributions with 4, 5, 6, and 7 degrees of freedom as these are the best fits in terms of sum of errors – as we will see just below – in normal and logarithmic scale respectively.



***Graph 19**: Zoom in left tail of distributions, logarithmic scale (Citigroup, April 7th 2014)*



***Graph 20**: Zoom in left tail of distributions, logarithmic scale (Citigroup, April 7th 2014)*



***Graph 21**: Zoom in left tail of distributions, logarithmic scale (Citigroup, April 7th 2014)*

III.4.B/ Results assessment

Squared errors and squared log-errors

We have started by computing errors between our estimated RN density and the calibrated Student distributions. In Table 8 we display the sum of squared errors and the sum of log squared errors for each distribution with respect to the average density. We observe that as we increase the degrees of freedom the SSE keeps decreasing up to 7 degrees. At 8 degrees of freedom, the SSE is higher than the previous distribution. In that sense, a Student distribution with 7 degrees of freedom seems to be an accurate choice to model the RN density.

Degrees of freedom	Sum of squared errors	Sum of log squared errors
2	11.8724	18,886.71
3	5.6819	8,001.47
4	2.7473	3,409.00
5	1.4162	1,459.44
6	0.8745	724.36
7	0.7312	578.47
8	0.7930	725.56

***Table 8:** Errors for each fitted non-standardized Student distribution*

We observe that as we increase the degrees of freedom the SSE keeps decreasing up to 7 degrees. At 8 degrees of freedom, the SSE is higher than the previous distribution. In that sense, a Student distribution with 7 degrees of freedom seems to be an accurate choice to model the RN density.

However, these squared errors are only based on the section of the distribution we have been able to estimate, namely from $29 to $57.5, hence these errors do not reflect how well our calibrated Student distribution captures tail events.

Option pricing

To address this shortcoming and have a further assessment on how accurate our calibrated Student risk neutral distributions are, we have backed American option prices from the distributions.

First, we are going to define more rigorously the RN distribution we have calibrated through the Student distribution, which means that we identify S_T with a random variable Student distributed, that we denote by S.

The corresponding RN distribution, given that the company has a non-zero probability of default, satisfies:

$$\mathbb{Q}(S_T = 0) \coloneqq Q_0 > 0.$$

i.e. the RN probability mass function is greater than 0 only at $S_T = 0$ and the above probability is equal to the probability that a non-standardized Student distribution takes negative values:

$$Q_0 = \mathbb{Q}(S \leq 0)$$

Note that:

$$\mathbb{Q} = Q_0\, \mathbb{Q}_{\mid S_T=0} + (1 - Q_0)\mathbb{Q}_{\mid S_T>0} = Q_0\, \mathbb{Q}_{\mid S_T=0} + (1 - Q_0)\mathbb{Q}_{\mid S>0}.$$

Denoting by $f_{\mathbb{Q}}$ be the density of the Student distribution, the risk-neutral distribution can thus be described as follows:

$$\mathbb{Q}\,(S_T \leq x) = \begin{cases} 0 & if\ x < 0 \\ Q_0 & if\ x = 0 \\ Q_0 + \int_0^x f_{\mathbb{Q}}(s)ds & if\ x > 0 \end{cases} = 1_{(x \geq 0)}\left(Q_0 + \int_0^x f_{\mathbb{Q}}(s)ds\right).$$

We can check that:

$\lim_{x\to-\infty} \mathbb{Q}\,(S \leq x) = 0$, and

$$\lim_{x\to+\infty} \mathbb{Q}\,(S \leq x) = Q_0 + \lim_{x\to+\infty} \int_0^x f_{\mathbb{Q}}(s)ds = Q_0 + (1 - Q_0) = 1.$$

Secondly, we must check that we can correctly derive American option prices from our risk neutral distributions.

Indeed, recall that our data refers to American option prices; if we want to use the pricing integral displayed in the introduction of the paper, which allows evaluating European options, we need to identify options which American price should be equal to the European price. We could also use an approximation technique that uses European option prices to estimate American option prices. We detail the choices available hereafter.

The price difference between European and American call options comes from dividend payments – see e.g. [27] – : if a stock does not distribute dividends, then the price of an American call option and a European call option are equal.

More generally, let us assume the underlying stock pays a dividend D on date t_D – in our Citigroup example, the value of the dividend is \$0.01 and the payment date May 1st, 2014 – and that both are known at valuation date $t = 0$. Then it is optimal to hold an American call option with strike k until its maturity date T if:

$$D \leq k\left(1 - e^{-r(T-t_D)}\right)$$

(see e.g. [19]). Therefore, if the above condition is met, the price of the American call is equal to its European counterpart.

In case this dividend condition is not met, it is still possible to easily approximate the American call price using Black's approximation. Letting $c(k,s_0,T)$ be the price of a European call written on a stock which current price is s_0, with strike k and maturity date T, Black's approximation to the price of an American call c^A is the following:

$$c^A(k, s_0, T) \approx \max\left(c(k, s_0 - De^{-rt_D}, T), c(k, s_0, t_D - 1)\right).$$

Black's approximation has been showed to work well most of the times. Its advantage over more sophisticated numerical methods – such as Barone-Adesi-Whaley or Roll-Geske-Whaley – is that it allows us to easily use our risk neutral distributions.

In our Citigroup dataset, the lowest strike call option has a strike equal to \$25. We use the quarterly dollar Libor rate r_L to approximate the bound on the dividend amount:

$$k\left(1-\frac{1}{1+r_L(T-t_D)}\right)=25\times\left(1-\frac{1}{1+0.227\%\times0.216}\right)\approx 0.0123$$

Given that the dividend paid out by Citigroup on May 1st, 2014, was 1 cent, and that the condition is met for the lowest strike, we conclude that the dividend condition is fulfilled for all strikes, thus we can price American calls with the European call pricing integral.

Unfortunately, the above results do not hold for put options.

Now, recall that the model-free price of a vanilla European call option is given by the formula:

$$c_k=e^{-rT}\mathbb{E}_{\mathbb{Q}}[(S_T-k)^+].$$

Given our modelling of the risk neutral distribution and assuming $k > 0$, our pricing integral becomes:

$$c_k=e^{-rT}\int_k^{\infty}(s-k)f_{\mathbb{Q}}(s)ds.$$

s represents the stock price, which goes up to infinity in the above integral. Therefore the price of the call option reflects tail events, in this case strong increases in the stock price. We can conclude that the closer our estimated option prices are to the empirical option prices, the better our calibrated Student distribution captures the market expectation of extreme price swings.

Given that we now have a series of estimated RN densities and using the quarterly Dollar Libor rate, we can approximate call prices for which we have empirical data with the following estimator:

$$\hat{c}_k=\frac{1}{1+r_LT}\sum_{i=1}^{N_\varepsilon}(s_i-k)\hat{f}_{\mathbb{Q}}(s_i)\Delta s_i.$$

where r_L is the Libor rate, T the time to maturity, k the strike and s_i a series of stock prices such that:

$$\forall i\in\mathbb{N},\Delta s_i=s_{i+1}-s_i=\Delta s$$

$$\forall i\in[1;N_\varepsilon]: s_i=k+i\Delta s.$$

N_ε verifies the following condition:

$$\forall i>N_\varepsilon: (s_i-k)\hat{f}_{\mathbb{Q}}(s_i)\Delta s<\varepsilon$$

where:

$$\varepsilon=10^{-9}.$$

The choice of 10^{-9} is motivated by computational considerations, as adding a further degree of accuracy – 10^{-10} – consumes much more time to complete the computations. It might be that N_ε is too low and therefore that we are not capturing tail events, however we have checked the values of s_i for which the estimator of the option price stops making computations and these goes from \$147.15 – for the \$57.5 strike and 8 degrees of freedom, equivalent to a stock return of 216.11% in 3 months; having a return higher than that has a RN probability of 2.3×10^{-8} under the 8 degree Student distribution – up to \$11,654.83 – for the \$29 strike with 2 degrees of freedom.

We also make the following test:

- We compute the integral for a Student distribution with 2 degrees of freedom and the \$29 strike – this is the larger integral – for $\varepsilon = 10^{-9}$ and $\varepsilon = 10^{-10}$; the first computation stops when s_i equals \$11,654.83 while the second one stops at s_i = \$36,736.15 – +215.20% – however the difference between the two approximations is just 0.00079 for an integral value which is very close to \$18 – approximately a 0.004% difference.
- We do the same for 8 degrees and \$57.5 strike, which corresponds to the smaller integral. The first computation stops at \$147.15 and the second one at \$181.83, the difference being 1.3×10^{-6} for an integral value that is close to \$0.083 – approximately a 0.002% difference.

We set Δs to be equal to \$0.01; again this is motivated by computational efficiency. For example, consider the call option with strike price of \$29; we have seen that the estimator includes stock prices up to \$11,654.83. With a \$0.01 step between stock prices, this amounts approximately to the following number of computations:

$$\frac{11{,}654.83-29}{0.01} \approx 1.17 \text{ million.}$$

Thus, if we used a much narrower step such as \$0.0001 to increase accuracy, we would end up making around 117 million computations to calculate the \$29 call price.

Finally, the estimator of the option price can thus be rewritten as:

$$\hat{c}_k = \frac{1}{1+r_L T}\sum_{i=1}^{N_\varepsilon} i(\Delta s)^2 \hat{f}_{\mathbb{Q}}(k+i\Delta s).$$

We use the RN density estimated through put prices. This allows us to test the robustness of our procedure, as we are using a density estimated through put values to derive call values. Given the negligible difference we found between both densities, we would expect to obtain good results.

We have displayed in Table 9 our estimated option prices for all seven Student distributions as well as the market price. In Table 10 estimated option prices are expressed as a percentage over the market price: $\frac{\hat{c}_k - c_k}{c_k}$.

Strike k	**Market c_k**	**Estimated $\hat{c}_k$**						
		η=2	**η=3**	**η=4**	**η=5**	**η=6**	**η=7**	**η=8**
\$29	*17.700*	17.9308	**17.6629**	17.6065	17.5864	17.5771	17.572	17.5689
\$30	*16.700*	16.953	**16.6755**	16.6153	16.5932	16.5827	16.5769	16.5733
\$31	*15.700*	15.9779	**15.6902**	15.6257	15.6016	15.5898	15.5832	15.5791
\$32	*14.750*	15.0057	**14.7073**	14.6385	14.612	14.599	14.5914	14.5867
\$33	*13.775*	14.0373	**13.7277**	13.6542	13.6253	13.6107	13.6023	13.5969
\$34	*12.825*	13.0732	**12.752**	12.6736	12.6422	12.6262	12.6168	12.6107
\$35	*11.725*	12.1145	11.7814	**11.698**	11.6641	11.6466	11.6363	11.6296
\$36	*10.800*	11.1624	**10.8172**	10.729	10.6925	10.6737	10.6626	10.6554
\$37	*9.825*	10.2187	**9.8615**	9.7685	9.7299	9.71	9.6984	9.6909
\$38	*9.075*	9.2854	**8.9168**	8.8197	8.7794	8.7589	8.7471	8.7397
\$39	*8.050*	8.3655	**7.9866**	7.8863	7.8452	7.8247	7.8132	7.8064
\$40	*7.050*	7.4633	**7.0757**	6.9736	6.9328	6.9132	6.9028	6.897
\$41	*6.125*	6.5841	6.1905	**6.0886**	6.0493	6.0316	6.0232	6.0191
\$42	*5.300*	5.7356	**5.3395**	5.2398	5.2037	5.1889	5.1831	5.1812
\$43	*4.575*	4.9281	**4.5333**	4.438	4.4062	4.3951	4.3923	4.393
\$44	*3.850*	4.1747	**3.7845**	3.6949	3.6679	3.6607	3.6611	3.6644

$45	*3.200*	3.4906	**3.1064**	3.0224	2.9996	2.9958	2.9988	3.0042
$46	*2.610*	2.8905	**2.5108**	2.4303	2.4102	2.4084	2.4129	2.4196
$47	*2.095*	2.3842	**2.0049**	1.9246	1.9047	1.903	1.9077	1.9144
$48	*1.655*	1.9725	**1.5891**	1.5057	1.4835	1.48	1.4833	1.4889
$49	*1.295*	1.6466	**1.2572**	1.1684	1.142	1.1354	1.1361	1.1397
$50	*0.995*	1.3921	**0.9977**	0.9032	0.872	0.8614	0.8591	0.8602
$52.5	*0.480*	0.9724	0.5807	**0.4784**	0.4384	0.4199	0.4107	0.4059
$55	*0.215*	0.7324	0.3623	0.2647	**0.2242**	0.2037	0.1919	0.1845
$57.5	*0.105*	0.5827	0.2418	0.1553	0.1198	**0.1014**	0.0906	0.0836

Table 9: *Estimated call prices for each fitted non-standardized Student distribution (s_0 = $46.55). The closest prices to the market are underlined in bold.*

Strike k	Estimated $\hat{c}_k$						
	η=2	η=3	η=4	η=5	η=6	η=7	η=8
$29	1.30%	**-0.21%**	-0.53%	-0.64%	-0.69%	-0.72%	-0.74%
$30	1.51%	**-0.15%**	-0.51%	-0.64%	-0.70%	-0.74%	-0.76%
$31	1.77%	**-0.06%**	-0.47%	-0.63%	-0.70%	-0.74%	-0.77%
$32	1.73%	**-0.29%**	-0.76%	-0.94%	-1.02%	-1.08%	-1.11%
$33	1.90%	**-0.34%**	-0.88%	-1.09%	-1.19%	-1.25%	-1.29%
$34	1.94%	**-0.57%**	-1.18%	-1.43%	-1.55%	-1.62%	-1.67%
$35	3.32%	0.48%	**-0.23%**	-0.52%	-0.67%	-0.76%	-0.81%
$36	3.36%	**0.16%**	-0.66%	-1.00%	-1.17%	-1.27%	-1.34%
$37	4.01%	**0.37%**	-0.58%	-0.97%	-1.17%	-1.29%	-1.36%
$38	2.32%	**-1.74%**	-2.81%	-3.26%	-3.48%	-3.61%	-3.69%
$39	3.92%	**-0.79%**	-2.03%	-2.54%	-2.80%	-2.94%	-3.03%
$40	5.86%	**0.36%**	-1.08%	-1.66%	-1.94%	-2.09%	-2.17%
$41	7.50%	1.07%	**-0.59%**	-1.24%	-1.52%	-1.66%	-1.73%
$42	8.22%	**0.75%**	-1.14%	-1.82%	-2.10%	-2.21%	-2.24%
$43	7.72%	**-0.91%**	-2.99%	-3.69%	-3.93%	-3.99%	-3.98%
$44	8.43%	**-1.70%**	-4.03%	-4.73%	-4.92%	-4.91%	-4.82%
$45	9.08%	**-2.92%**	-5.55%	-6.26%	-6.38%	-6.29%	-6.12%
$46	10.75%	**-3.80%**	-6.89%	-7.66%	-7.72%	-7.55%	-7.30%
$47	13.80%	**-4.30%**	-8.13%	-9.08%	-9.16%	-8.94%	-8.62%
$48	19.18%	**-3.98%**	-9.02%	-10.36%	-10.57%	-10.37%	-10.04%
$49	27.15%	**-2.92%**	-9.78%	-11.81%	-12.32%	-12.27%	-11.99%
$50	39.91%	**0.27%**	-9.23%	-12.36%	-13.43%	-13.66%	-13.55%
$52.5	102.58%	20.98%	**-0.33%**	-8.67%	-12.52%	-14.44%	-15.44%
$55	240.65%	68.51%	23.12%	**4.28%**	-5.26%	-10.74%	-14.19%
$57.5	454.95%	130.29%	47.90%	14.10%	**-3.43%**	-13.71%	-20.38%

Table 10: *Estimated call prices as a percentage over market price for each fitted non-standardized Student distribution. The minimal percentage is indicated in bold.*

A first and quick way to assess which distribution yields the best result is to look at how close the cheapest call estimated and market prices – with strike $57.5 – are. We immediately see that the Student distribution with 6 degrees of freedom is the best performer in that sense. However, it also looks like the Student distribution with 3 degrees of freedom performs better when pricing options with greater moneyness, 3 being closer to the empirical tail index of stock prices.

However, a more precise analysis is required. We have to look at how separated estimated and market prices are globally. To do so, we cannot compute squared errors but instead percentage errors: indeed, a particular distribution could yield a sum of squared errors higher than another distribution, but from a price perspective what matters is how far apart prices are in percentage terms. Thus, for each distribution and for each strike we have computed statistics for the following measures – absolute error and relative error:

$$e_{rel} = \left|\frac{\hat{c}_k}{c_k} - 1\right| ; \qquad e_{abs} = \frac{\hat{c}_k}{c_k} - 1.$$

It is important to look at the maximum of the above statistics for each distribution. Indeed, from a practitioner perspective we want our estimated prices to be on average close to market prices on a relative basis, but we also want to minimize the likelihood of having outlier estimations. We also show the number of times the distribution is the most accurate in the absolute error table.

Results are showed in Tables 11 and 12.

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8
Average	39.32%	9.92%	5.62%	4.45%	4.41%	5.15%	5.57%
St. Dev.	1.00	0.29	0.10	0.04	0.04	0.05	0.06
Maximum	454.95%	130.29%	47.90%	14.10%	13.43%	14.44%	20.38%
Minimum	1.30%	0.06%	0.23%	0.52%	0.67%	0.72%	0.74%

Table 11: *Relative error statistics for estimated call prices*

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8
Average	39.32%	7.94%	0.07%	-2.98%	-4.41%	-5.15%	5.57%
St. Dev.	1.00	0.29	0.12	0.05	0.04	0.05	0.06
Maximum	454.95%	130.29%	47.90%	14.10%	-0.67%	-0.72%	-0.74%
Minimum	1.30%	-4.30%	-9.78%	-12.36%	-13.43%	-14.44%	-20.38%
Best dist.	0	20	3	1	1	0	0

Table 12: *Absolute error statistics for estimated call prices*

If we look at the relative error, the best distribution according to the statistics we have defined has 6 degrees of freedom, which is consistent with the computation of SSE, as the Student distribution with $\eta = 6$ was the one with the second lowest errors: its average relative difference is 4.41%. If we look at the maximum relative difference, this 6 degree distribution has also the best performance. Overall, Student distributions with 4, 5, 6, 7 and 8 degrees of freedom perform relatively well, with an average relative difference smaller than 6%. However, if we look at each option price independently, we notice that the Student distribution with 3 degrees also performs very well: it is the best pricing distribution for 20 strikes, so 80% of our dataset. Its average results are distorted by the most out-of-the-money calls – $52.5, $55 and $57.5 – where it yields an average relative difference of 73.26%. However, if we exclude these options, it achieves an impressive 1.28% average difference.

If we now look at absolute error statistics, we notice that distributions with 2, 3 and 4 degrees of freedom tend to overestimate the price of the call, while the rest tends to underestimate it. This feature is natural because, as we increase the degrees of freedom of the Student distribution, the implied probability of extreme upward price movements' decreases, therefore decreasing the value of the call's optionality.

We have computed the same statistics but excluding the three most OTM options – strikes at \$52.5, \$55 and \$57.5. These figures confirm what we highlighted above, namely that by excluding the OTM options we observe that the distribution with 3 degrees of freedom yields the best results.

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8
Average	8.40%	1.28%	3.14%	3.83%	4.05%	4.09%	4.05%
St. Dev.	0.10	0.01	0.03	0.04	0.04	0.04	0.04
Maximum	39.91%	4.30%	9.78%	12.36%	13.43%	13.66%	13.55%
Minimum	1.30%	0.06%	0.23%	0.52%	0.67%	0.72%	0.74%

Table 13: *Relative error statistics for estimated call prices excluding the three most OTM options*

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8
Average	8.40%	-0.96%	-3.14%	-3.83%	-4.05%	-4.09%	-4.05%
St. Dev.	0.10	0.02	0.03	0.04	0.04	0.04	0.04
Maximum	39.91%	1.07%	0.23%	-0.52%	-0.67%	-0.72%	-0.74%
Minimum	1.30%	-4.30%	-9.78%	-12.36%	-13.43%	-13.66%	-13.55%
Best dist.	0	20	2	0	0	0	0

Table 14: *Absolute error statistics for estimated call prices excluding the three most OTM options*

One final check we can make on the quality of our Student distributions is to price options that we did not include in the estimation process – those for which the bid-offer spread was wider than 35% or that did not have an implied volatility associated to them. As these options tend to be relatively illiquid, they might not accurately reflect market expectations and thus the estimation procedure will probably yield prices than deviate significantly from market observed prices. It is nonetheless interesting to test the limits of our procedure and this final test should shed further light on how reliable our RN density estimation method is. We have priced calls we did not use in the estimation of the density. There are 6 of them in total. Results are showed in Table 15.

Strike k	Market c_k	Estimated $\hat{c}_k$						
		η=2	η=3	η=4	η=5	η=6	η=7	η=8
\$25	*21.6*	21.8609	21.6268	**21.5832**	21.5692	21.5634	21.5604	21.5588
\$26	*20.575*	20.876	20.6342	20.5878	**20.5725**	20.5659	20.5625	20.5606
\$27	*19.55*	19.8926	19.6426	19.593	19.5763	19.5689	19.5651	**19.5628**
\$28	*18.575*	18.9107	18.6521	18.5992	18.5809	**18.5726**	18.5681	18.5655
\$60	*0.075*	0.482	0.1708	0.0967	**0.0674**	0.0527	0.0441	0.0386
\$65	*0.05*	0.3567	0.0967	**0.0434**	0.0249	0.0164	0.0119	0.0092

Table 15: *Estimated call prices for testing options*
(i.e. excluded from the estimation/calibration phase)

Strike k	Estimated $\hat{c}_k$						
	η=2	η=3	η=4	η=5	η=6	η=7	η=8
\$25	1.21%	0.12%	**-0.08%**	-0.14%	-0.17%	-0.18%	-0.19%
\$26	1.46%	0.29%	0.06%	**-0.01%**	-0.04%	-0.06%	-0.07%
\$27	1.75%	0.47%	0.22%	0.13%	0.10%	0.08%	**0.07%**
\$28	1.81%	0.42%	0.13%	0.03%	**-0.01%**	-0.04%	-0.05%
\$60	542.67%	127.73%	28.93%	**-10.13%**	-29.73%	-41.20%	-48.53%
\$65	613.40%	93.40%	**-13.20%**	-50.20%	-67.20%	-76.20%	-81.60%

Table 16: *Estimated call prices for testing options as a percentage over market price (i.e. excluded from the estimation/calibration phase)*

As done above, we have computed a series of statistics to assess the accuracy of our method. Results are showed in Table 17 and 18.

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8
Average	193.72%	37.07%	7.10%	10.11%	16.21%	19.63%	21.75%
St. Dev.	2.99	0.58	0.12	0.20	0.28	0.32	0.35
Maximum	613.40%	127.73%	28.93%	50.20%	67.20%	76.20%	81.60%
Minimum	1.21%	0.12%	0.06%	0.01%	0.01%	0.04%	0.05%

Table 17: *Relative error statistics for test call prices*

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8
Average	193.72%	37.07%	2.68%	-10.05%	-16.18%	-19.60%	-21.73%
St. Dev.	2.99	0.58	0.14	0.20	0.28	0.32	0.35
Maximum	613.40%	127.73%	28.93%	0.13%	0.10%	0.08%	0.07%
Minimum	1.21%	0.12%	-13.20%	-50.20%	-67.20%	-76.20%	-81.60%
Best dist.	0	0	2	2	1	0	1

Table 18: *Absolute error statistics for test call prices*

The first observation is that the accuracy has greatly diminished compared to the pricing of call options that were used to estimate densities. Moreover, the best performing distribution seems now to be a non-standardized Student with 4 degrees of freedom.

Again, it is worth analyzing the impact of OTM options on accuracy: as we increase further the number of degrees, the estimated price decreases drastically, particularly for the \$65 call. More precisely, if we look at distributions with 3 and 4 degrees of freedom, we see that the estimated price transitions from 193% of the market value to 87%.

If we now look only at ITM calls, we can see that the estimation seems much more accurate than for the OTM calls – strikes at \$60 and \$65. We have thus calculated the same statistics but excluding these two options. Results are showed in Tables 19 and 20.

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8
Average	1.56%	0.33%	0.12%	0.08%	0.08%	0.09%	0.09%
St. Dev.	~ 0.0	~ 0.0	~ 0.0	~ 0.0	~ 0.0	~ 0.0	~ 0.0
Maximum	1.81%	0.47%	0.22%	0.14%	0.17%	0.18%	0.19%
Minimum	1.21%	0.12%	0.06%	0.01%	0.01%	0.04%	0.05%

Table 19: *Relative error statistics for test call prices excluding OTM options*

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8
Average	1.56%	0.33%	0.08%	~ 0.0%	-0.03%	-0.05%	-0.06%
St. Dev.	~ 0.0	~ 0.0	~ 0.0	~ 0.0	~ 0.0	~ 0.0	~ 0.0
Maximum	1.81%	0.47%	0.22%	0.13%	0.10%	0.08%	0.07%
Minimum	1.21%	0.12%	-0.08%	-0.14%	-0.17%	-0.18%	-0.19%
Best dist.	0	0	1	1	1	0	1

Table 20: *Absolute error statistics for test call prices excluding OTM options*

Our intuition is confirmed as results have improved a lot. Distribution with 5 degrees of freedom is the best performer, particularly if we look at absolute errors. Nonetheless these last results have to be considered cautiously, as this last sample of estimated option prices contains 4 points only – strikes at $25, $26, $27 and $28 respectively.

The decrease in accuracy when pricing test options – *i.e.* options not included in the estimation and calibration phases – seems to be totally explained by OTM options. This phenomenon could be attributed to the tails of Student distribution, which we might fail to accurately model due to the lack of data: this phenomenon highlights the importance of the tails of distributions and indicates that a bespoken modelling of these might significantly improve our results.

However, related to this aspect, it is also worth mentioning that the two OTM test options – $60 and $65 strikes – have a bid-offer spread of 100% and 800% respectively. In this context, it is far more challenging for our estimated densities to capture market expectations, as such large spreads convey that trading on those price ranges – $60-$65 – is extremely scarce and option values might not convey any meaningful market expectation. Meanwhile, the ITM call test options just tested have an average bid-offer spread equal to 18.55%. Trading is more buoyant in those ranges and option prices incorporate market expectations better, thus our densities perform better.

Therefore, one could reasonably ask whether there is a fundamental shortcoming in our distribution tail modelling or instead there is too little liquidity to properly assess the ability of our Student distribution to capture implied tail events. We not delve further into this question as the purpose of this paper is not to develop an option pricing technique.

	ITM calibration calls	*OTM calibration calls*	*ITM test calls*	*OTM test calls*
Average bid-offer spread	11.04%	3.96%	18.55%	450.00%

Table 21: *Average market bid-offer spreads for different groups of options*

Finally, we have also used the test set to assess the impact of the stopping criterion ε on the integral approximation: we have computed test option prices by setting $\varepsilon = 10^{-10}$ to analyze whether there is any meaningful accuracy gain. In the next table, we have displayed the percentage difference between this 2nd group of estimated prices and the 1st one, where $\varepsilon = 10^{-9}$.

Strike k	η=2	η=3	η=4	η=5	η=6	η=7	η=8
$25	0.0037%	0.0005%	~0.0%	0.0005%	~0.0%	~0.0%	~0.0%
$26	0.0038%	0.0005%	~0.0%	~0.0%	~0.0%	~0.0%	~0.0%
$27	0.0040%	~0.0%	0.0005%	~0.0%	~0.0%	~0.0%	~0.0%
$28	0.0042%	~0.0%	~0.0%	~0.0%	~0.0%	~0.0%	~0.0%
$60	0.1660%	0.0585%	~0.0%	~0.0%	~0.0%	~0.0%	~0.0%
$65	0.1962%	0.1034%	0.2304%	~0.0%	~0.0%	~0.0%	~0.0%

Table 22: *Test options estimated prices' sensitivity to ε*

As expected, the accuracy gains affect mostly the distribution $\eta = 2$. As we increase the number of degrees, the distribution tails get thinner and the accuracy gain coming from decreasing the value of ε decreases rapidly. Nonetheless, overall these accuracy gains remain extremely low.

Further key statistics relating to our option pricing results can be found on the appendix. We end this discussion by considering two Tables. In the first one, Table 23, we show how the percentage difference with respect to ATM options: these correspond to the $47 and $46 call.

Strike k	η=2	η=3	η=4	η=5	η=6	η=7	η=8
$46	10.75%	**-3.80%**	-6.89%	-7.66%	-7.72%	-7.55%	-7.30%
$47	13.80%	**-4.30%**	-8.13%	-9.08%	-9.16%	-8.94%	-8.62%
Average	12.28%	**-4.05%**	-7.51%	-8.37%	-8.44%	-8.25%	-7.96%
Abs. avrg.	12.28%	**4.05%**	7.51%	8.37%	8.44%	8.25%	7.96%

Table 23: *Estimated option prices for ATM options as a percentage over market price*

In Table 24, we show the same figures with respect to OTM options.

Strike k	η=2	η=3	η=4	η=5	η=6	η=7	η=8
$47	13.80%	**-4.30%**	-8.13%	-9.08%	-9.16%	-8.94%	-8.62%
$48	19.18%	**-3.98%**	-9.02%	-10.36%	-10.57%	-10.37%	-10.04%
$49	27.15%	**-2.92%**	-9.78%	-11.81%	-12.32%	-12.27%	-11.99%
$50	39.91%	**0.27%**	-9.23%	-12.36%	-13.43%	-13.66%	-13.55%
$52.5	454.95%	130.29%	47.90%	14.10%	**-3.43%**	-13.71%	-20.38%
$55	240.65%	68.51%	23.12%	**4.28%**	-5.26%	-10.74%	-14.19%
$57.5	102.58%	20.98%	**-0.33%**	-8.67%	-12.52%	-14.44%	-15.44%
$60	542.67%	127.73%	28.93%	**-10.13%**	-29.73%	-41.20%	-48.53%
$65	613.40%	93.40%	**-13.20%**	-50.20%	-67.20%	-76.20%	-81.60%
Average	228.25%	47.78%	**5.58%**	-10.47%	-18.18%	-22.39%	-24.93%
Abs. avrg.	228.25%	50.26%	**16.63%**	14.55%	18.18%	22.39%	24.93%
Best dist.	0	**4**	2	2	1	0	0

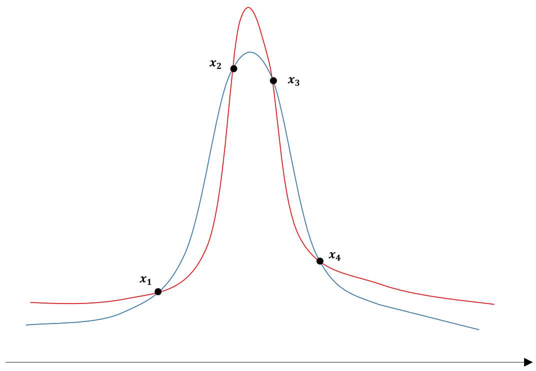
Table 24: *Estimated option prices for OTM options as a percentage over market price*

These last results show that, in terms of pricing, choosing the best distribution depends significantly on your purpose: if you are mostly interested by highly liquid options, the best distribution is the one that has only 3 degrees; however, if your focus is on extreme tail events and how those are reflected on market prices, the Student distribution with 4 or 5 degrees of freedom seems a much better fit.

Paradoxically, this result might be considered as counterintuitive: as we get further OTM and option prices become more sensible to extreme events, we observe that the number of degrees of the optimal Student distribution increases, which means that it assigns lower probabilities to tail returns. However, developments such as the 2008-2009 financial crisis might tend to indicate that market agents underestimate the likelihood of extreme events, thus we could have expected the number of degrees of the best distribution to decrease as we got further into OTM territory.

Nevertheless, this last remark has to be nuanced:

- First and foremost we have to bear in mind that we are working with RN distributions, not real world probabilities;
- Let St_1 and St_2 be two Student distributions with roughly the same location and scale parameters, such St_2 has more degrees of freedom than St_1. As it can observed in distributions graphs throughout this paper, St_1 is going to cross St_2 at 4 points x_1, x_2, x_3 and x_4, such that St_1's density will be above St_2's density for values lower than x_1, higher than x_4 or between x_2 and x_3.



  Thus, we can see that the probability assigned to extreme events depends on what we call an extreme event, *i.e.* the threshold that determines we are in “tail territory”. Indeed, if we consider that upward tail events begin above x_4, then unambiguously distribution St_1 assigns a higher probability to these scenarios; however, if we consider that extreme scenarios begin at some point between x_3 and x_4, then distribution St_2 might assign a higher probability to some of these scenarios. In this sense, our results might in fact not be counterintuitive.

Finally, we undertake the same process but for put options this time. Although, as we have explained above, there are no bounds on dividends or computational shortcuts that allow us to obtain exact put prices, we have decided to make these computations because Citigroup’s dividend is very low:

- The dividend yield for 2014 is 0.086% ;
- The dividend amount to be paid on May 1st, 2014 – 1 cent – represents just 0.040% of the lowest strike – $25.

We thus believe that the distortion in the put price coming from ignoring the dividend payment should be negligible.

The price estimation methodology is equivalent to the one implemented for call options. Results can be found in the appendix.

III.4.C/ Analysis of relevant events and their associated probabilities

Here, we analyze our distributional results by considering relevant events, such as for example company default, and computing their probabilities.

Risk neutral default probabilities

First, it is worth commenting that by modelling our risk neutral density with a non-standardized Student distribution, we allow the stock price to take negative values. However, we explained previously that this can be interpreted as the RN probability of default of the underlying bank. Let *X* be a random variable distributed according to a non-standardized Student distribution and t_S a random variable distributed according to a Student distribution with *x* degrees of freedom. Then the probability that *X* could take negative values is equal to:

$$\mathbb{P}(X \leq 0) = \mathbb{P}\left(\frac{X-\mu}{\sigma} \leq -\frac{\mu}{\sigma}\right) = \mathbb{P}\left(t_S \leq -\frac{\mu}{\sigma}\right).$$

The probabilities returned by R for each parametrization of the Student distribution are:

Degrees of freedom	$\mathbb{P}(X \leq 0)$
2	0.3074%
3	0.0688%
4	0.0200%
5	0.0069%
6	0.0027%
7	0.0012%
8	0.0005%

Table 25: *3-month risk neutral probabilities of default for fitted non-standardized Student distributions*

We can interpret these probabilities as the 3-month risk neutral probability of bankruptcy for Citigroup. If we now assume that quarterly default is an independent random variable, these probabilities can be annualized by remarking that: $1 - PD_A = (1 - PD_T)^4 84914641$, where PD_A is the annual probability of default and PD_T is the quarterly probability of default. It comes:

$$PD_A = 1 - (1 - PD_T)^4.$$

In Table 26, we display annualized default probabilities. We have also added the closest corresponding S&P's credit rating notation – S&P calculates a one-year Average Cumulative Default Rate for each credit rating category based on historical default data from 1981; the 2014 figures can be consulted in [29]. Citigroup LT local issuer rating in April 2014 was "A-", corresponding to a historical default rate of 0.08%.

Degrees of freedom	Annualized probabilities	S&P rating
2	1.2239%	BB- (1.09%)
3	0.2749%	BBB- (0.30%)
4	0.0800%	A- (0.08%)
5	0.0276%	AA- (0.03%)
6	0.0108%	AA (0.02%)
7	0.0048%	AA+, AAA (0.00%)
8	0.0020%	AA+, AAA (0.00%)

Table 26: *One-year risk neutral probabilities of default for fitted non-standardized Student distributions assuming independence*

Another option would be to assume dependence. In such a case, we use a result due to Groenendijk et al. [15] which states that, provided that the 2nd moment of the distribution converges – which is the case in our setting – the following scaling law applies:

$$\sigma_A = \sqrt{\frac{1}{TtM}} \times \sigma_T = \sqrt{\frac{1}{0{,}282}} \times \sigma_T \approx \sqrt{3{,}55} \times \sigma_T.$$

σ_A is the annual volatility – the standard deviation in our Student distribution – while σ_T is the quarterly volatility and TtM the time to maturity. Annual probabilities are displayed in Table 27:

Degrees of freedom	Annualized probabilities	S&P rating
2	1.0650%	BB- (1.09%)
3	0.4306%	BB+ (0.40%)
4	0.2203%	BBB (0.19%)
5	0.1303%	BBB+ (0.13%)
6	0.0850%	A- (0.08%)
7	0.0595%	A+ (0.06%)
8	0.0439%	AA- (0.03%)

Table 27: *One-year risk neutral probabilities of default for fitted non-standardized Student distributions assuming non independence*

As expected, an increasing number of degrees of freedom is associated with a decrease in the default probability, given that by nature the bank's default is a tail event.

We can compare these probabilities to those computed by S&P for guidance – always bearing in mind that S&P calculates real probabilities as opposed to risk neutral ones.

We use Table 28 to compare S&P's and our own probabilities. Figures correspond to the ratio between the two:

$$Ratio = \frac{PD_{Student}}{PD_{S\&P}}$$

Degrees of freedom	Annualized probabilities relative to S&P's one-year default probability (independence hyp)	Annualized probabilities relative to S&P's one-year default probability (non-independence hyp)
2	15.30	13.31
3	3.44	5.38
4	1.00	2.75
5	0.35	1.63
6	0.14	1.06
7	0.06	0.74
8	0.03	0.55

Table 28: *Comparison of fitted default rates (independence hypothesis in the middle column and non independence hypothesis in the right one) and S&P default rates*

Risk neutral strikes' probabilities

In this paragraph we show the RN probabilities associated to each quoted option strike for each calibrated Student distribution. Letting t_S be a Student-distributed random variable, they have been obtained by using the following equality:

$$\mathbb{Q}\,(S_T \leq k) = \mathbb{Q}\left(t_S \leq \frac{k-\mu}{\sigma}\right)$$

Strike k	Probability $\mathbb{Q}(S_T \leq k)$						
	η=2	η=3	η=4	η=5	η=6	η=7	η=8
$25	1.39%	0.63%	0.36%	0.24%	0.17%	0.13%	0.10%
$26	1.52%	0.72%	0.43%	0.29%	0.21%	0.16%	0.13%
$27	1.67%	0.83%	**0.51%**	0.35%	0.27%	0.21%	0.18%
$28	1.84%	0.95%	0.61%	0.44%	0.34%	0.28%	0.24%
$29	2.05%	1.10%	0.73%	**0.55%**	**0.44%**	0.37%	0.32%
$30	2.28%	1.29%	0.89%	0.69%	0.57%	**0.49%**	**0.44%**
$31	2.56%	1.52%	1.09%	0.87%	0.74%	0.65%	0.59%
$32	2.90%	1.80%	1.35%	1.11%	0.97%	0.88%	0.81%
$33	3.30%	2.16%	1.68%	1.43%	1.28%	1.18%	1.12%
$34	3.78%	2.60%	2.11%	1.85%	1.70%	1.61%	1.54%
$35	4.38%	3.17%	2.67%	2.42%	2.28%	2.19%	2.13%
$36	5.12%	3.91%	3.42%	3.19%	3.06%	2.99%	2.95%
$37	6.04%	4.87%	4.42%	4.22%	4.13%	4.10%	4.08%
$38	7.23%	6.13%	5.75%	5.62%	5.59%	5.60%	5.62%
$39	8.76%	7.79%	7.54%	7.50%	7.55%	7.62%	7.70%
$40	10.76%	10.02%	9.93%	10.02%	10.16%	10.30%	10.44%
$41	13.42%	12.98%	13.09%	13.33%	13.57%	13.80%	14.00%
$42	16.95%	16.89%	17.23%	17.61%	17.94%	18.24%	18.48%
$43	21.66%	21.98%	22.50%	22.97%	23.36%	23.68%	23.95%
$44	27.81%	28.39%	28.98%	29.46%	29.84%	30.14%	30.38%
$45	35.53%	36.11%	36.59%	36.96%	37.23%	37.45%	37.63%
$46	44.55%	44.82%	45.02%	45.17%	45.28%	45.37%	45.44%
$47	54.14%	53.94%	53.78%	53.67%	53.58%	53.52%	53.47%
$48	63.30%	62.75%	62.29%	61.95%	61.69%	61.49%	61.33%
$49	71.22%	70.62%	70.04%	69.57%	69.20%	68.91%	68.67%
$50	77.59%	77.22%	76.68%	76.20%	75.81%	75.49%	75.22%
$52.5	87.64%	88.20%	88.16%	87.98%	87.77%	87.57%	87.39%
$55	92.59%	93.67%	94.03%	94.15%	94.18%	94.16%	94.13%
$57.5	95.17%	96.38%	96.87%	97.12%	97.25%	97.33%	97.37%
$60	96.64%	97.79%	98.27%	98.52%	98.67%	98.76%	98.83%
$65	98.13%	99.03%	99.38%	99.55%	99.65%	99.71%	99.75%

Table 30: *Risk neutral strikes' probabilities*

Risk neutral quantiles

We have also looked at some relevant quantiles for our fitted distributions and compared them to historical quantiles derived from Citigroup's stock price history. The question that naturally arises is how should risk neutral and real world probabilities compare? On one hand, assuming there is no dividend – *i.e.* $d = 0$ – recall that under the risk neutral distribution we should have:

$$\mathbb{E}_{\mathbb{Q}}[S_T] = s_0 e^{rT}.$$

That is, the expected risk neutral return of the stock is equal to the risk free rate. On the other hand, financial theory postulates that a risky investment should reward investors with an expected return μ higher than the risk-free rate ($\mu > r$):

$$\mathbb{E}_{\mathbb{P}}[S_T] = s_0 e^{\mu T}$$

Thus, if we designate the real world probability measure by $\mathbb{P}$, the following inequality holds:

$$\mathbb{E}_{\mathbb{P}}[S_T] > \mathbb{E}_{\mathbb{Q}}[S_T].$$

Although we have $\mathbb{E}_{\mathbb{P}}[S_T] = \int_0^\infty \mathbb{P}[S_T > x]dx$, it does not help to compare the distribution of S_T under $\mathbb{P}$ and $\mathbb{Q}$, respectively. In an empirical study on options on the FTSE100 by Humphreys and Noss (2012) (see [20]), it is shown that measuring with $\mathbb{Q}$ overestimates the losses and underestimates the profits (see Chart 26 p. 22 in [20]):

$$(S)\begin{cases} \mathbb{P}(S_T > x) \geq \mathbb{Q}(S_T > x), \forall x > s_0 \\ \mathbb{P}(S_T \leq x) \leq \mathbb{Q}(S_T \leq x), \forall x < s_0 \end{cases}$$

i.e. the real world measure assigns a higher likelihood to prices increases than the risk neutral measure. When looking at the first two moments, the authors obtain the same comparison for the mean as we have, $\mathbb{E}_{\mathbb{P}}[X] > \mathbb{E}_{\mathbb{Q}}[X]$, and the variance satisfies the reverse inequality, $var_{\mathbb{P}}(X) \leq var_{\mathbb{Q}}(X)$ (see Charts 28 & 29 p. 23 in [20]). In another paper by Giordano and Siciliano (see [14]), this result comparing the P&L distributions under those two probability measures is confirmed and proved when considering a simplified binomial framework (see [14], Boxes 1 & 4, and Figure 1). If an investor is risk-averse – *i.e.* his utility function is concave – then the real world probability $\mathbb{P}(S_T > s_0)$ is shown to be higher than the risk neutral probability $\mathbb{Q}(S_T > s_0)$. Picking up the same reasoning, it can be easily shown that the inverse is true if the investor is risk prone – *i.e.* his utility function is convex. Thus, comparing risk neutral and real world quantiles could also allow us to state whether, on an aggregate basis, investors on a particular asset are either risk averse or risk prone.

From the above relations given in (S), we can deduce, by definition of quantiles, that for any given threshold α and time T,

$$q_\alpha^{\mathbb{P}}(S_T) \geq q_\alpha^{\mathbb{Q}}(S_T),$$

which means that we expect a higher Value-at-Risk of the price S_T under the real-world probability $\mathbb{P}$ than the one obtained under the risk-neutral probability $\mathbb{Q}$, for the same threshold.

Now, when looking at returns, we should get a lower VaR of the return under the risk neutral probability than under the real-world one (since the signs of the returns are negative).

In the following tables, we have summarized risk neutral quantiles for different Student distributions, when looking at quantiles below the median, focusing on negative ones. They have been obtained by using the following equality:

$$\alpha = \mathbb{P}(t_S \leq q_\alpha) = \mathbb{P}(X \leq \mu + \sigma q_\alpha),$$

where the non standardized rv X has been introduced previously.

In Table 31, quantiles are expressed as quarterly returns without annualization:

$$\frac{(\mu + \sigma q_\alpha) - s_0}{s_0}$$

with $s_0 = \$46.55$.

As the returns are for a quarter, they can be relatively high in absolute value and therefore logarithmic returns are not suitable.

Quantile	η=2	η=3	η=4	η=5	η=6	η=7	η=8
0.01%	-100.00%	-100.00%	-100.00%	-92.57%	-79.24%	-71.48%	-66.47%
0.05%	-100.00%	-100.00%	-78.88%	-65.69%	-58.83%	-54.72%	-52.01%
0.1%	-100.00%	-88.06%	-65.71%	-56.36%	-51.41%	-48.41%	-46.43%
0.5%	-78.17%	-50.34%	-42.16%	-38.55%	-36.59%	-35.39%	-34.60%
1%	-54.84%	-39.12%	-34.31%	-32.16%	-31.01%	-30.32%	-29.86%
5%	-22.97%	-20.26%	-19.50%	-19.24%	-19.16%	-19.14%	-19.16%
10%	-14.82%	-14.09%	-14.01%	-14.08%	-14.18%	-14.29%	-14.38%
25%	-6.39%	-6.56%	-6.75%	-6.91%	-7.05%	-7.16%	-7.25%
50%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%

Table 31: *Relevant quantiles for fitted non-standardized Student distributions*

In order to put into perspective the quarterly returns showed on the previous tables, we compute historical quantiles for Citigroup's stock returns. To do so, we first retrieve historical prices adjusted for splits and dividends (*source: Yahoo Finance*). Then, we have computed returns, which are approximately quarterly, from April 9th, 1984, to April 7th, 2014.

Time to maturity for our panel of options is equal to 0.282, but based on a full year – *i.e.* taking into account weekend days. If we make the standard assumption that a year has 252 working days, multiplying this figure by 0.282 we obtain a time to maturity of 71 business days. Therefore to compute Citigroup's quarterly stock returns we consider a 70-day lag:

$$R_t^{Trimestral} = \frac{S_t - S_{t-70}}{S_{t-70}}$$

We compute quantiles for the following stock price quarterly returns time series, all ending on April 7th, 2014: 2 years (502 observations), 5 years (1,258 observations), 10 years (2,516 observations), 15 years (3,774 observations) and 30 years (7,563 observations). Note that we compute those quarterly returns by a window moving every day, which means using overlapping daily observations. Hence, the extreme quantiles beyond 1% should be taken with caution as they are computed from strongly dependent observations.

Results are shown on Table 32. We notice there that quantiles do not converge with the number of observations, except for the extreme quantile (0.01%). They are quite unstable. The first quartile is negative while the median becomes positive, reflecting the risk premium attributed to equities. For the extreme quantile, we move from no data (2 years) to 7 overlapping observations (30 years), thus the saturation in the high quantiles, as well as the drastic change in values, contrary to the median.

Quantile	2 years	5 years	10 years	15 years	30 years
0.01%	-31.38%	-59.66%	-86.19%	-86.06%	-85.64%
0.05%	-31.23%	-59.16%	-85.07%	-84.34%	-82.56%
0.1%	-31.04%	-57.72%	-83.82%	-82.57%	-80.55%
0.5%	-29.36%	-45.37%	-78.22%	-77.01%	-72.68%
1%	-28.31%	-40.33%	-75.00%	-72.71%	-59.71%
5%	-22.09%	-29.27%	-40.15%	-33.40%	-31.09%
10%	-16.85%	-22.88%	-29.16%	-24.44%	-19.59%
25%	-2.91%	-9.26%	-11.58%	-8.67%	-5.93%
50%	8.59%	3.78%	-0.53%	1.14%	4.76%

Table 32: *Historical quantiles for Citigroup's stock quarterly return*

For better comparison, we have added to Table 33 the historical quantiles from the 30 years series:

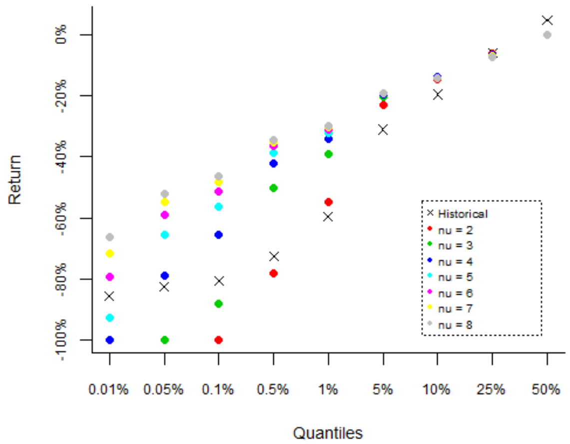
Quantile	30yr	η=2	η=3	η=4	η=5	η=6	η=7	η=8
0.01%	*-85.64%*	-100.00%	-100.00%	-100.00%	-92.57%	**-79.24%**	-71.48%	-66.47%
0.05%	*-82.56%*	-100.00%	-100.00%	**-78.88%**	-65.69%	-58.83%	-54.72%	-52.01%
0.1%	*-80.55%*	-100.00%	**-88.06%**	-65.71%	-56.36%	-51.41%	-48.41%	-46.43%
0.5%	*-72.68%*	**-78.17%**	-50.34%	-42.16%	-38.55%	-36.59%	-35.39%	-34.60%
1%	*-59.71%*	**-54.84%**	-39.12%	-34.31%	-32.16%	-31.01%	-30.32%	-29.86%
5%	*-31.09%*	**-22.97%**	-20.26%	-19.50%	-19.24%	-19.16%	-19.14%	-19.16%
10%	*-19.59%*	**-14.82%**	-14.09%	-14.01%	-14.08%	-14.18%	-14.29%	-14.38%
25%	*-5.93%*	**-6.39%**	-6.56%	-6.75%	-6.91%	-7.05%	-7.16%	-7.25%
50%	*4.76%*	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%

Table 33: *Relevant quantiles for fitted non-standardized Student distributions together with historical quantiles for 30-year returns. Bold values are the closest to the historical ones.*

We have finally computed the sum of squared errors between Student distributions' quantiles and historical quantiles for the 30-year series:

Degrees of freedom	Sum of squared errors
2	0.1054
3	0.1660
4	0.2205
5	0.3036
6	0.3775
7	0.4458
8	0.5000

Table 34: *Sum of squared errors for each fitted non-standardized Student distribution*



Graph 22*: 30-year historical and distributional quantiles (Citigroup, April 7th 2014)*

In Graph 22, we present graphically 30-years historical quantiles as well as quantiles for the Student models. We observe that from the 1% quantile up to the first quartile, historical VaR are below the RN ones, contrary to what we claim in equations (S). We also notice that beyond 1%, the historical quantiles are leveling off, illustrating most probably the lack of observations. Empirical estimations of the tails have shown [5] that the degree of freedom of the Student-t distribution describing best the empirical distribution of financial returns should be around 3 or 4.

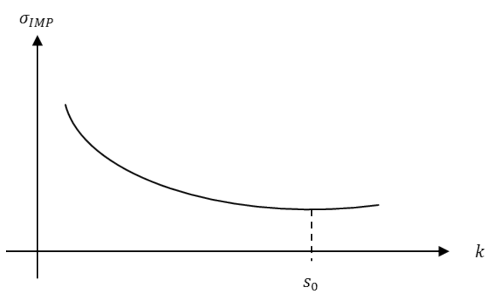
Our conclusion is therefore that we are observing the opposite of what we were expecting. A few explanations could be put forward:

1. First, as we explained, the comparison of risk neutral and real world probability measures could be interpreted in terms of markets agents' risk preferences. In this case, it could be argued that on aggregate basis investors in Citigroup's stock are risk prone.
2. Another important interpretation is that investors fail to accurately price extreme events (beyond the first quartile).

The second explanation is in our view the most plausible one, as does the havoc experienced by financial markets during the 2008-2009 crisis show. The mispricing of extreme events by market participants can be explained by various reasons related, among other things, to demand and supply dynamics:

- Holding deep OTM options can be viewed as an expensive trade, given that, in case the option is held to maturity, the initial investment is totally lost if the extreme event has not materialized. Therefore a lot of market participants could be deterred from entering this kind of trades, which should result in a structurally weak demand and low market prices. Low market prices in turn translate into a low implied likelihood of extreme events.
- It is relatively well known that a common practice in financial markets is to sell OTM puts, cashing in the premium while expecting the put will not be exercised; this strategy should drive up OTM puts supply and down OTM puts prices, thus lowering the risk neutral likelihood of extreme downfalls implied by these market prices.
- More generally, the repetition throughout financial markets history of crashes and crisis has lead experts and analysts to state that market participants have a short memory – see for example *This time is different – Eight centuries of financial folly* from C. M. Reinhart and K. Rogoff [24]. Whether this is due to pure negligence or to market dynamics that penalize investors that do not "ride the wave" of bull markets, the fact is that financial markets seem subject periodically to episodes of what R. Shiller has dubbed "*irrational exuberance*" [26], during which financial prices experience sharp revaluations that led investors to underestimate the likelihood of crashes.

Although our results might indicate that market participants improperly price the likelihood of extreme events, there is evidence that option prices reflect to some extent adverse tail scenarios: since the stock crash of 1987, equity option markets [17] show a "volatility skew" – also known as "volatility smile", for example in FX option markets. As observed in Graphs 4, 5 and 6 – for convenience we have also schematically represented this skew in the figure below – the market's implied volatility is a concave function of the strike price, with a minimum at $k = s_0$. In the case of the equity option market, the slope of the curve for $k < s_0$, tends to be higher than the slope when $k > s_0$, which is the reason why we speak of a skew instead of a smile.



Option pricing models always represent the option's price as monotonically increasing in volatility, hence the above pattern indicates that – on a volatility-basis – OTM options are more expensive than ATM options. Our analysis shows that the pricing of extreme events is still insufficient – the volatility skew should be sharper.

IV – Final results and concluding remarks

We choose to concentrate on the SIFIs to explore if the market includes a systemic risk component in the pricing of options of those institutions. This component should translate in fat tails for the implied risk neutral distribution from option prices. Moreover, the option market for these large institutions is very liquid and should be quite representative of market perception. We developed our method on Citigroup option prices, then apply the same statistical procedure to three more American financial institutions belonging to the list of global SIFIs: Bank of America Merrill Lynch, Goldman Sachs and Morgan Stanley. Results for the other three banks – as well as for Citigroup in the case where selected option prices have a bid-ask spread equal to or lower than 10% – can be found in the appendix. Hereafter we display and discuss the main results for the four banks.

	Citigroup*	**Bank of America Merrill Lynch***	**Goldman Sachs***	**Morgan Stanley***
Best Student distribution in terms of log SE w.r.t average RN density	$\eta = 7$ *(see Table 8)*	$\eta = 4$ *(see Table A.20)*	$\eta = 5$ *(see Table A.43)*	$\eta = 8$ *(see Table A.66)*
Closest Student distribution in terms of SE for quantile 1.0%	$\eta = 2$ *(see Table 33)*	$\eta = 2$ *(see Table A.36)*	$\eta = 2$ *(see Table A.59)*	$\eta = 3$ *(see Table A.83)*
Closest Student distribution in terms of SE for quantile 5.0%	$\eta = 2$ *(id)*	$\eta = 2$ *(id)*	$\eta = 7$ *(id)*	$\eta = 2$ *(id)*
Closest Student distribution in terms of SE for quantile 25.0%	$\eta = 2$ *(id)*	$\eta = 4$ *(id)*	$\eta = 7$ *(id)*	$\eta = 2$ *(id)*
Best Student distribution in terms of option pricing** *(calibration & test options)*	$\eta = 3$ *(see Tables 9 & 10)*	$\eta = 3$ *(see Table A.23)*	$\eta = 2$ *(see Table A.46)*	$\eta = 4$ *(see Table A.69)*
Best Student distribution in terms of ATM option pricing*** *(calibration & test options)*	$\eta = 3$ *(see Table 23)*	$\eta = 3$ *(see Table A.29)*	$\eta = 7$ *(see Table A.52)*	$\eta = 5$ *(see Table A.75)*
Best Student distribution in terms of OTM option pricing** *(calibration & test options)*	$\eta = 3$ *(see Table 24)*	$\eta = 6$ *(see Table A.30)*	$\eta = 7$ *(see Table A.53)*	$\eta = 7$ *(see Table A.76)*

* Only call options are considered.

** The best distribution is selected according to the number of times it is the most accurate distribution.

*** The best distribution is selected according to the average absolute error.

Table 35: *Best fit for non-standardized Student distributions according to various criteria*

	Citigroup (30 years)	**Bank of America Merrill Lynch** (25 years)	**Goldman Sachs** (14,75 years)	**Morgan Stanley** (25 years)
Empirical longest years 1% quantile	-59.71%	-55,64%	-50,40%	-50,22%
1% quantile computed with the fitted Student with η = 3 distribution	-39,12%	-41,55%	-32,51%	-48,18%
Difference between empirical (RW) and fitted RN quantiles	-20,59%	-14,09%	-17,89%	-2,04%
Closest Student	η = 2	η = 2	η = 2	η = 3

Table 36: *Comparison of the RW and the RN 1%-quantiles for four SIFIs*

When picking the best distribution in the option pricing category, for "all call options" and "OTM call options", we have made the choice to rank distributions according to the number of accurate valuations, instead of the average pricing error, as in some cases an extremely large valuation error on a single option might distort an otherwise excellent accuracy on the rest of options.

We now discuss the main takeaways from Tables 35 & 36 :

- We notice that generally the best distributions in terms of sum of log squared errors – with respect to the estimated average RN density – have more degrees of freedom than those whose quantiles are the closest to empirical quantiles – line 1 against lines 2, 3 and 4 of Table 35.

  As it has already been the case, this result is inconsistent with our understanding of risk neutral and real world probabilities: under the RN world financial assets yield on average the risk free rate, thus downwards events should have a greater probability mass under the RN distribution than under the real world one; we might then expect the RN distribution to have fatter tails than what we observe in the prices, which translates into higher likelihood of extreme events (except for Goldman Sachs from 5% quantile onwards). When comparing RN quantiles to real world quantiles, we would expect the closest Student distributions to have more degrees of freedom than the optimal (in terms of log squared errors) distribution, not less, as observed in our study.
- Looking at Table 35, we see the difficulty of accurately estimate the tails. We find scattered responses and sometimes contradictory when using various estimation criteria. We have discussed that when talking about Citigroup. The same holds for the other studied banks. We expect also that using other dates for the option prices would also lead to similar results.
- Regarding option pricing, on a general basis, Student distributions with few degrees of freedom seem to perform better for the ATM options, while it is the reverse for OTM ones. This seems paradoxical as the OTM options should be more sensitive to fat tails.
- For ATM option pricing, we observe a low number of degrees of freedom for Citigroup and Bank of America (η = 3) as well as Morgan Stanley (η = 5) but a relatively high

number for Goldman Sachs (η = 7). Nevertheless, note that, for Goldman Sachs, the distribution η = 2 also performs very well – average absolute error equal to 2.00% against 1.18% for η = 7.

- On the contrary, for OTM pricing, distributions with a high number of degrees of freedom perform better, the exception being Citigroup.

  This phenomenon echoes what we already observe for Citigroup's deep OTM call options: OTM options are better priced by Student distributions with more degrees of freedom than those of optimal Student distributions for pricing ATM options. This phenomenon might be considered paradoxical with regards to financial markets history; although we reasoned that the paradox might be solved by considering the definition of an extreme event.

  - We make a last remark concerning OTM pricing results. It is only for Citigroup that a strictly correct pricing approach has been undertaken, as for Bank of America Merrill Lynch, Goldman Sachs and Morgan Stanley we decided to ignore the effect of dividend payments on the ground that their amount were small and thus their impact on American option prices should be negligible. These last results might indicate that the impact has not been in fact negligible; however, we have not been able to explain why ignoring dividends would lead to having an optimal pricing distribution with more degrees of freedom. If anything, pricing an American call as a European one should lead to an undervaluation of the option, therefore we would expect the best distribution to be one with very few degrees of freedom.

To conclude, despite the difficulties of assessing the tail, it seems clear that the assumption that the RN distribution should overestimate the tail risk, is not verified in our study. As mentioned previously, there are two possible explanations for that: either investors are risk prone as far as the tails are concerned, or the market does not price the extreme risks. Both are not contradictory: a risk prone investor will not price correctly the extreme risks.

References

[1] Black, F. and Scholes, M. (1973), "The pricing of Options and Corporate Liabilities", *The Journal of Political Economy*, Vol. 81, No. 3, 637-654

[2] Breeden, D.T., and Litzenberger, R. H. (1978), "Prices of state-contingent claims implicit in option prices", *The Journal of Business*, Vol. 51, No. 4, 621-651

[3] Dacorogna, M. M. (2014), "The Price of Being a SIFI", *ESSEC Conference on Extreme Events in Finance*, Abbaye de Royaumont (France), December 15-17, 2014

[4] Dacorogna, M. and Busse, M. (2016), "The price of Being a Systematically Important Financial Institution (SIFI)", *Available at SSRN (http://ssrn.com/abstract=2803113)*

[5] Dacorogna, M., Pictet, 0. , Müller, U.A. and De Vries, C.G. (2001), "Extremal forex returns in extremely large data sets", *Extremes*, Vol. 4(2), 105-127.

[6] Delbaen, F. and Schachermayer, W. (1994) "The fundamental theorem of asset pricing", *Mathematische Annalen*, Vol. 300, Issue 1, 463-520

[7] Delecroix, M. (2006), *Statistiques non paramétriques*, Notes de cours ISUP (UPMC)

[8] Eberlein, E. and Madan, D. B. (2011), "Unbounded Liabilities, Reserve Capital Requirements and the Taxpayer Put Option", *Quantitative Finance*, Vol. 12, Issue 5, 709-724

[9] Everitt, B. S. and Hothorn, T. (2006), *A Handbook of Statistical Analyses Using R*. Chapman and Hall

[10] Faraway, J. J. (2006), *Extending the Linear Model with R*. Chapman & Hall

[11] Figlewski, S. (2008), "Estimating the Implied Risk Neutral Density for the U.S. Market Portfolio", *Volatility and Time Series Econometrics: Essay in Honor of Robert F. Engle*. Oxford University Press

[12] Financial Stability Board (2014), "2014 update of list of global systemically important banks (G-SIBs)", Financial Stability Board

[13] Geraty, M. (2000), "How to Calculate the Odds of a Change in the Fed Funds Rate", Bianco Research L.L.C.

[14] Giordano, L. and Siciliano, G. (2013), "Real-world and risk-neutral probabilities in the regulation on the transparency of structured products", Commissione Nazionale per le Società e la Borsa

[15] Groenendijk P. A., Lucas, A., and de Vries, C. G. (1996), "Stochastic processes, nonnormal innovations, and the use of scaling ratios", *Proceedings of the third International Conference on Forecasting Financial Markets*, London March 27-29, 1996, 1, 1–38.

[16] Hait, D.J. (2001), "OptionMetrics: Dividend Forecasts, Option Pricing Models and Implied Volatility Calculations", OptionMetrics

[17] Hansen, B.E. (2009), *Lecture Notes on Nonparametrics*, Univ. Wisconsin, available online (www.ssc.wisc.edu/~bhansen/718/NonPaarametrics1.pdf)

[18] Heston, S.L. (1993), "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options", *The Review of Financial Studies*, Vol. 6(2), 327-343

[19] Hull, J.C. (2014), *Options, Futures and Other Derivatives*. Pearson

[20] Humphreys, V. and Noss J. (2012), Estimating probability distributions of future asset prices: empirical transformations from option-implied risk-neutral to real-world density functions, Bank of England, Working Paper.

[21] Körner, K., "Deutsche Bank Research: Extracting implied default probabilities from CDS spreads", online tool

[22] Lai, W.-N. (2010), "Comparison of Methods to Estimate Option Implied Risk Neutral Densities", *Quantitative Finance*, Vol. 14, Issue 10, 1839-1855

[23] Lander, J. P. (2014), *R for Everyone, Advanced Analytics and Graphics*. Addison-Wesley Professional

[24] Reinhart, C.M. and Rogoff, K. (2009), *This time is different – Eight centuries of financial folly*. Princeton University Press

[25] Ross, S. (2015), "The Recovery Theorem", Journal of Finance, Vol. 70 (2), 615-648

[26] Shiller, R. (2015), *Irrational Exuberance*. Princeton University Press (3rd Ed.), Princeton, New Jersey

[27] Shreve, S. (2004), *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer Finance

[28] Silverman, B.(1998), *Density Estimation for Statistics and Data Analysis*. Chapman&Hall

[29] Standard & Poor's Ratings Services (2014), "2014 Annual Global Corporate Default Study And Rating Transitions", table 26, Standard & Poor's Ratings Services

[30] Tian, Y. S. (2010), "Extracting Risk-Neutral Density and its Moments from American Option Prices", *Journal of Derivatives*, Vol. 10, No. 3, 17-34

Appendix

A – Additional Citigroup results

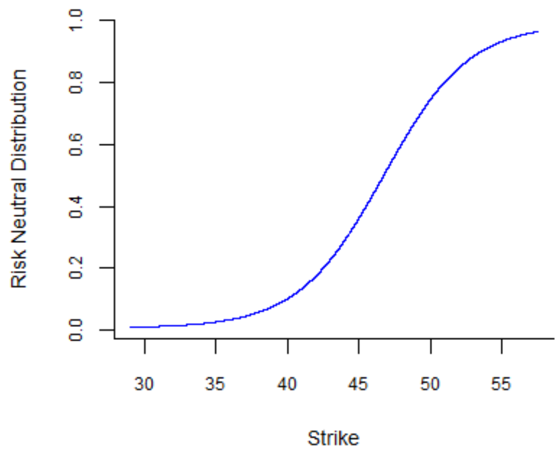
Hereafter we present some additional graphs and tables relating to Citigroup results that we do not deem to be as critical as to be included in the main text.

We also present some additional results relating to option put pricing to evaluate the accuracy of our Student distribution calibration procedure.

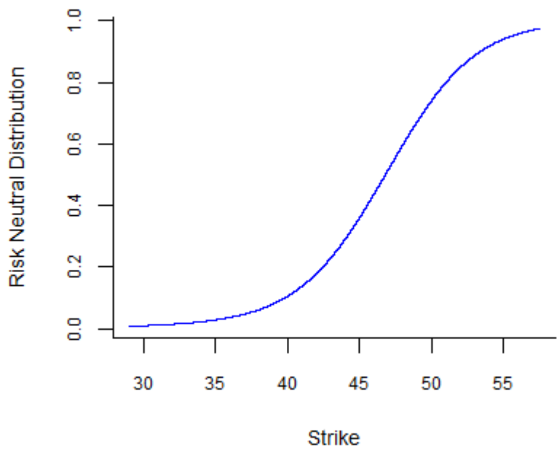
For each one of the appendix's subsections and subsubsections, we have written in parenthesis the corresponding section and subsection numbering in the main text.

A.1/ Risk Neutral Distribution of stock prices: technical methodology and results (III)

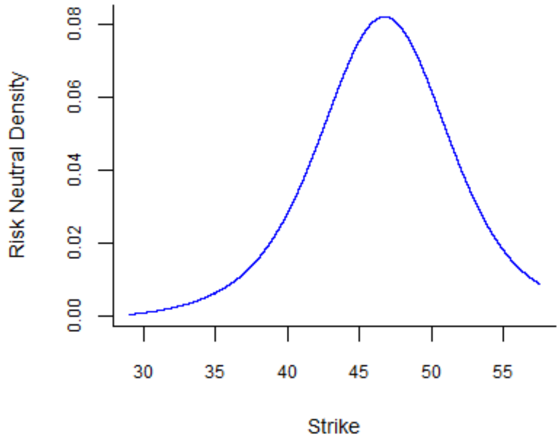
A.1.1/ Estimating the risk neutral probability (III.1)



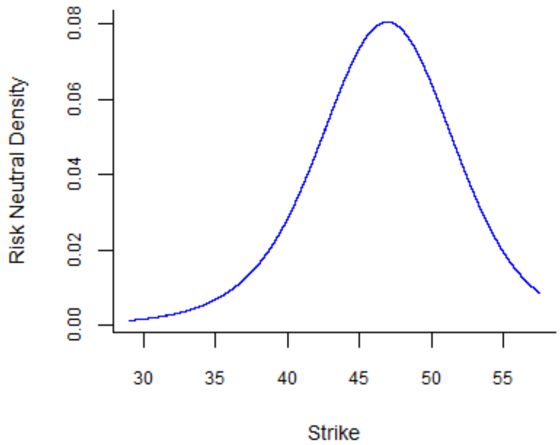
Graph A.1*: Risk neutral distribution for the polynomial interpolation method (Citigroup, April 7th 2014)*



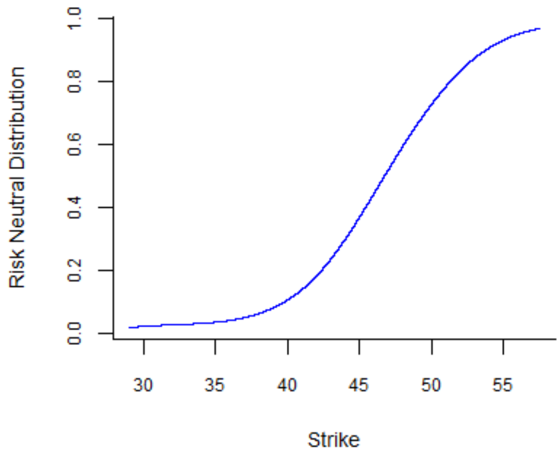
Graph A.2*: Risk neutral distribution for the spline smoothing method (Citigroup, April 7th 2014)*



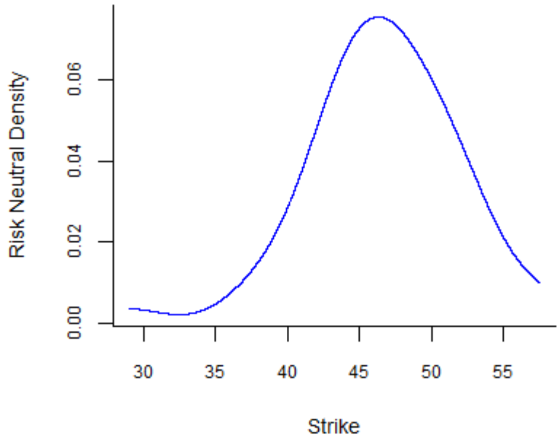
Graph A.3: *Risk Neutral Density for the polynomial interpolation method (Citigroup, April 7th 2014)*



Graph A.4: *Risk Neutral Density for the spline smoothing method (Citigroup, April 7th 2014)*



Graph A.5: *Amended risk neutral distribution for Nadaraya-Watson estimator h = 9 (Citigroup, April 7th 2014)*



Graph A.6*: Amended Risk Neutral Density for Nadaraya-Watson estimator h = 9 (Citigroup, April 7th 2014)*

A.1.2/ Fitting a known parametric probability distribution (III.4)

Option pricing for call options (Citigroup)

The below tables display additional key statistics for all call options – calibration calls, test calls.

Statistic	**η=2**	**η=3**	**η=4**	**η=5**	**η=6**	**η=7**	**η=8**
Average	69.20%	15.17%	5.90%	5.55%	6.70%	7.96%	8.70%
St. Dev.	1.63	0.37	0.10	0.09	0.13	0.15	0.17
Maximum	613.40%	130.29%	47.90%	50.20%	67.20%	76.20%	81.60%
Minimum	1.21%	0.06%	0.06%	0.01%	0.01%	0.04%	0.05%

Table A.1: *Relative error statistics for estimated option prices for all options*

Statistic	**η=2**	**η=3**	**η=4**	**η=5**	**η=6**	**η=7**	**η=8**
Average	4.13%	0.82%	1.52%	1.86%	2.00%	2.05%	2.06%
St. Dev.	0.03	0.01	0.02	0.02	0.02	0.02	0.02
Maximum	10.75%	3.80%	6.89%	7.66%	7.72%	7.55%	7.30%
Minimum	1.21%	0.06%	0.06%	0.01%	0.01%	0.04%	0.05%

Table A.2: *Relative error statistics for estimated ITM option prices for all options*

Statistic	**η=2**	**η=3**	**η=4**	**η=5**	**η=6**	**η=7**	**η=8**
Average	7.34%	1.13%	2.68%	3.26%	3.44%	3.47%	3.44%
St. Dev.	0.09	0.01	0.03	0.04	0.04	0.04	0.04
Maximum	39.91%	4.30%	9.78%	12.36%	13.43%	13.66%	13.55%
Minimum	1.21%	0.06%	0.06%	0.01%	0.01%	0.04%	0.05%

Table A.3: *Relative error statistics for estimated option prices for all options ex. the 5 most OTM*

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8
Average	69.20%	13.58%	0.57%	-4.35%	-6.69%	-7.95%	-8.69%
St. Dev.	1.63	0.37	0.12	0.10	0.13	0.15	0.17
Maximum	613.40%	130.29%	47.90%	14.10%	0.10%	0.08%	0.07%
Minimum	1.21%	-4.30%	-13.20%	-50.20%	-67.20%	-76.20%	-81.60%
Best dist.	0	20	5	3	2	0	1

Table A.4: *Absolute error statistics for estimated option prices for all options*

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8
Average	4.13%	-0.41%	-1.48%	-1.85%	-1.99%	-2.04%	-2.05%
St. Dev.	0.03	0.01	0.02	0.02	0.02	0.02	0.02
Maximum	10.75%	1.07%	0.22%	0.13%	0.10%	0.08%	0.07%
Minimum	1.21%	-3.80%	-6.89%	-7.66%	-7.72%	-7.55%	-7.30%
Best dist.	0	16	3	1	1	0	1

Table A.5: *Absolute error statistics for estimated ITM option prices for all options*

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8
Average	7.34%	-0.77%	-2.64%	-3.24%	-3.43%	-3.47%	-3.44%
St. Dev.	0.09	0.02	0.03	0.04	0.04	0.04	0.04
Maximum	39.91%	1.07%	0.22%	0.13%	0.10%	0.08%	0.07%
Minimum	1.21%	-4.30%	-9.78%	-12.36%	-13.43%	-13.66%	-13.55%
Best dist.	0	20	3	1	1	0	1

Table A.6: *Absolute error statistics for estimated option prices for all options excluding the 5 most OTM*

Option pricing for put options (Citigroup)

Strike k	Market p_k	Estimated $\hat{p}_k$						
		η=2	η=3	η=4	η=5	η=6	η=7	η=8
$37	*0.165*	0.514	0.283	0.203	0.167	0.147	0.136	0.129
$38	*0.220*	0.580	0.338	0.254	0.216	0.196	0.184	0.177
$39	*0.295*	0.659	0.407	0.320	0.281	0.261	0.250	0.243
$40	*0.395*	0.756	0.495	0.406	0.368	0.349	0.338	0.333
$41	*0.530*	0.876	0.609	0.521	0.484	0.466	0.458	0.454
$42	*0.705*	1.027	0.758	0.671	0.637	0.623	0.617	0.616
$43	*0.925*	1.219	0.951	0.869	0.839	0.829	0.826	0.827
$44	*1.200*	1.465	1.202	1.125	1.100	1.094	1.094	1.098
$45	*1.545*	1.780	1.523	1.452	1.431	1.428	1.431	1.437
$46	*1.955*	2.180	1.927	1.859	1.841	1.840	1.845	1.851
$47	*2.445*	2.673	2.420	2.353	2.335	2.334	2.339	2.346
$48	*3.020*	3.260	3.004	2.933	2.913	2.910	2.914	2.920
$49	*3.650*	3.934	3.671	3.595	3.571	3.565	3.566	3.570
$50	*4.350*	4.679	4.411	4.329	4.300	4.291	4.288	4.289
$52.5	*6.350*	6.757	6.492	6.403	6.365	6.347	6.338	6.334

Table A.7: *Estimated put prices for each fitted non-standardized Student distribution (S_0 = $46.55)*

Strike k	Estimated $\widehat{p}_k$						
	η=2	η=3	η=4	η=5	η=6	η=7	η=8
$37	211.21%	71.52%	23.03%	**1.03%**	-10.67%	-17.64%	-22.12%
$38	163.45%	53.45%	15.23%	**-2.00%**	-11.09%	-16.41%	-19.73%
$39	123.42%	37.90%	8.31%	**-4.85%**	-11.63%	-15.42%	-17.76%
$40	91.44%	25.37%	**2.84%**	-6.94%	-11.75%	-14.33%	-15.77%
$41	65.36%	14.98%	**-1.79%**	-8.75%	-12.00%	-13.57%	-14.32%
$42	45.72%	7.49%	**-4.81%**	-9.60%	-11.62%	-12.43%	-12.68%
$43	31.79%	**2.80%**	-6.10%	-9.29%	-10.42%	-10.70%	-10.62%
$44	22.08%	**0.12%**	-6.26%	-8.32%	-8.87%	-8.82%	-8.54%
$45	15.23%	**-1.44%**	-6.04%	-7.36%	-7.57%	-7.37%	-7.01%
$46	11.49%	**-1.46%**	-4.91%	-5.82%	-5.88%	-5.64%	-5.30%
$47	9.31%	**-1.03%**	-3.78%	-4.49%	-4.54%	-4.34%	-4.07%
$48	7.96%	**-0.54%**	-2.88%	-3.54%	-3.63%	-3.52%	-3.33%
$49	7.77%	**0.58%**	-1.50%	-2.16%	-2.33%	-2.30%	-2.20%
$50	7.55%	1.40%	**-0.48%**	-1.14%	-1.37%	-1.42%	-1.39%
$52.5	6.41%	2.24%	0.83%	0.24%	**-0.04%**	-0.18%	-0.26%

Table A.8: *Estimated put prices as a percentage over market price for each fitted non-standardized Student distribution*

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8
Average	54.68%	14.82%	5.92%	5.03%	7.56%	8.94%	9.67%
St. Dev.	0.65	0.22	0.06	0.03	0.04	0.06	0.07
Maximum	211.21%	71.52%	23.03%	9.60%	12.00%	17.64%	22.12%
Minimum	6.41%	0.12%	0.48%	0.24%	0.04%	0.18%	0.26%

Table A.9: *Relative error statistics for estimated call prices*

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8
Average	54.68%	14.23%	0.78%	-4.87%	-7.56%	-8.94%	-9.67%
St. Dev.	0.65	0.23	0.09	0.04	0.04	0.06	0.07
Maximum	211.21%	71.52%	23.03%	1.03%	-0.04%	-0.18%	-0.26%
Minimum	6.41%	-1.46%	-6.26%	-9.60%	-12.00%	-17.64%	-22.12%
Best dist.	0	7	4	3	1	0	0

Table A.10: *Absolute error statistics for estimated call prices*

We have also priced some out-of-sample puts; we have decided to ignore some prices from our panel. Indeed, for strikes below $34 most options violate non-arbitrage conditions, as can be seen in Table 2. Let k_1 and k_2 be two different strikes, then by no-arbitrage we must have:

$$k_2 > k_1 \Rightarrow p_{k_2} > p_{k_1}$$

Yet we can observe in the aforementioned table that market prices for put options at strikes $26, $27, $28, $29, $30, $32 and $33 break this condition. In such a case, modelling results might seem inaccurate and distorted, and it is wiser to not take into account these prices when assessing our method.

Strike k	Market p_k	Estimated $\widehat{p}_k$						
		η=2	η=3	η=4	η=5	η=6	η=7	η=8
\$34	*0.070*	0.370	0.175	0.110	0.081	0.066	0.056	0.050
\$35	*0.080*	0.411	0.204	0.134	0.102	0.085	0.075	0.068
\$36	*0.110*	0.458	0.239	0.164	0.130	0.112	0.101	0.094
\$55	*8.380*	9.016	8.772	8.688	8.650	8.630	8.618	8.611
\$57.5	*10.780*	11.364	11.150	11.077	11.044	11.026	11.015	11.008
\$60	*13.180*	13.762	13.578	13.516	13.490	13.475	13.467	13.462
\$65	*18.430*	18.634	18.500	18.460	18.444	18.436	18.432	18.429

Table A.11: *Estimated put prices for testing options*
(i.e. excluded from the estimation/calibration phase)

Strike k	Estimated $\widehat{p}_k$						
	η=2	η=3	η=4	η=5	η=6	η=7	η=8
\$34	428.57%	150.57%	57.14%	15.57%	**-6.43%**	-19.71%	-28.29%
\$35	413.25%	155.13%	67.25%	27.63%	6.50%	**-6.13%**	-14.50%
\$36	316.27%	117.55%	49.09%	18.18%	**1.55%**	-8.36%	-14.91%
\$55	7.59%	4.68%	3.67%	3.22%	**2.98%**	2.84%	2.75%
\$57.5	5.42%	3.43%	2.75%	2.44%	2.28%	2.18%	**2.12%**
\$60	4.42%	3.02%	2.55%	2.35%	2.24%	2.18%	**2.14%**
\$65	1.10%	0.38%	0.16%	0.07%	0.03%	**0.01%**	**-0.01%**

Table A.12: *Estimated put prices for testing options as a percentage over market price*
(i.e. excluded from the estimation/calibration phase)

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8
Average	168.09%	62.11%	26.09%	9.92%	3.14%	5.92%	9.24%
St. Dev.	2.07	0.75	0.30	0.11	0.02	0.07	0.10
Maximum	428.57%	155.12%	67.25%	27.62%	6.50%	19.71%	28.29%
Minimum	1.10%	0.38%	0.16%	0.07%	0.03%	0.01%	0.01%

Table A.13: *Relative error statistics for test put prices*

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8
Average	168.09%	62.11%	26.09%	9.92%	1.31%	-3.86%	-7.24%
St. Dev.	2.07	0.75	0.30	0.11	0.04	0.08	0.12
Maximum	428.57%	155.12%	67.25%	27.62%	6.50%	2.84%	2.75%
Minimum	1.10%	0.38%	0.16%	0.07%	-6.43%	-19.71%	-28.29%
Best dist.	0	0	0	0	3	2	3

Table A.14: *Absolute error statistics for test put prices*

B – Computations for other banks

This second appendix compiles additional results for risk neutral distribution estimation for three other banking corporations: Bank of America Merrill Lynch, Goldman Sachs and Morgan Stanley. Moreover, we have also included a section in which we test the sensitivity of our estimation procedure to the choice of option data by restricting the choice of Citigroup's options we pick.

B.1/ Bank of America Merrill Lynch

In this section, we have undertaken the same procedures as previously but applied to options written on Bank of America Merrill Lynch (BAML) stock. Data is from April 1st, 2014. Bank of America Merrill Lynch shares closed at $17.34 on that day.

B.1.1/ Data

Given that compared to Citigroup's Bank of America Merrill Lynch's dataset is more limited, we have kept all option mid prices for which we have an implied volatility value for our estimation procedure. However, except for the $22 call and the $13 put, all other options have a bid-offer spread narrower than 35%, so our choice of data seems reasonable.

Calls			
Strike	*Bid-offer spread*	*Mid*	*Implied volatility*
10	1.37%	7.35	-
11	1.59%	6.35	-
12	**1.87%**	**5.4**	**39.99%**
13	**2.30%**	**4.4**	**32.56%**
14	**1.47%**	**3.425**	**27.77%**
15	**0.40%**	**2.525**	**26.14%**
16	**0.59%**	**1.705**	**24.25%**
17	**0.96%**	**1.045**	**23.20%**
18	**1.75%**	**0.575**	**22.57%**
19	**3.57%**	**0.285**	**22.24%**
20	**7.69%**	**0.135**	**22.40%**
21	**16.67%**	**0.065**	**22.99%**
22	**50.00%**	**0.025**	**22.73%**
23	100.00%	0.015	-
24	-	0.005	-
25	-	0.015	-
26	-	0.01	-

Puts			
Strike	*Bid-offer spread*	*Mid*	*Implied volatility*
10	-	0.015	-
11	-	0.015	-
12	200.00%	0.02	-
13	**66.67%**	**0.04**	**30.29%**
14	**12.50%**	**0.085**	**27.94%**
15	**5.88%**	**0.175**	**25.73%**
16	**2.86%**	**0.355**	**23.95%**
17	**1.43%**	**0.705**	**23.23%**
18	**0.81%**	**1.235**	**22.60%**
19	**0.52%**	**1.945**	**22.27%**
20	**0.36%**	**2.795**	**22.44%**
21	**2.70%**	**3.75**	**24.82%**
22	**2.15%**	**4.7**	**24.72%**
23	**1.77%**	**5.7**	**28.38%**
24	**1.50%**	**6.7**	**31.81%**
25	**1.31%**	**7.7**	**35.03%**
26	1.74%	8.675	-

Tables A.15 and A.16*: Option prices data for BAML on April 1st, for a 109 days maturity (2014)*

B.1.2/ Risk neutral density estimation

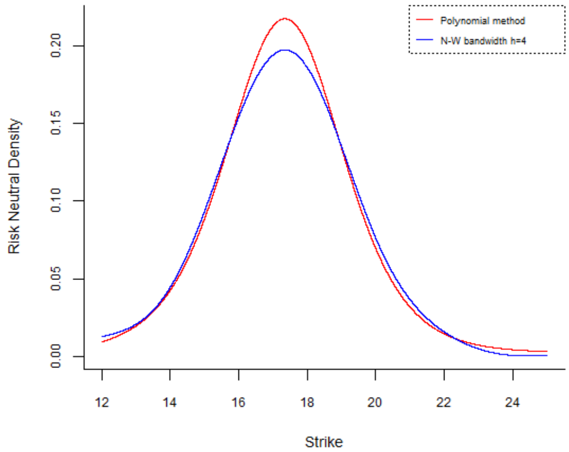
With respect to Citigroup, polynomial interpolation has been undertaken with a 3 degree polynomial in the case of Bank of America Merrill Lynch. Furthermore, given the structure of BAML option prices data the spline smoothing methods yielded unsatisfactory results – extremely bumpy interpolation, negative density. We decided to exclude it from the estimation procedure.

Method	Polynomial interpolation	Spline smoothing	Nadaraya-Watson estimator: h = h*	Nadaraya-Watson estimator: h = 4
SSE for implied volatility estimation	0.0008747	-	0.0018329	0.0049147

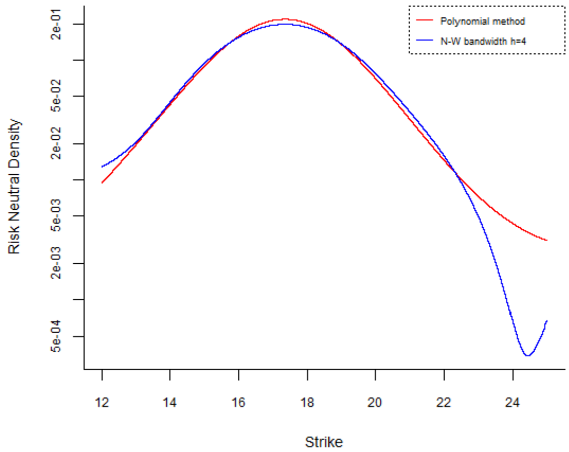
Table A.17: *Sum of Squared Errors for all implied volatility estimation methods*

Method	Polynomial interpolation	Spline smoothing	Nadaraya-Watson estimator: h = h*	Nadaraya-Watson estimator: h = 4
Density integral estimate	99.07%	-	96.56%	97.00%

Table A.18: *Density integral estimates for polynomial, spline and Nadaraya-Watson methods*



Graph A.7: *Estimated RNDs for BAML by polynomial and Nadaraya-Watson interpolation*



Graph A.8: *Estimated RNDs for BAML, logarithmic scale*

In addition, we observe in the above graph that the RND estimated through the Nadaraya-Watson method has some undesirable properties, as it was the case for Citigroup. We thus decided to exclude it from the rest of the procedure so in the case of Bank of America we have not computed an average RND but simply retained the RND obtained by polynomial interpolation.

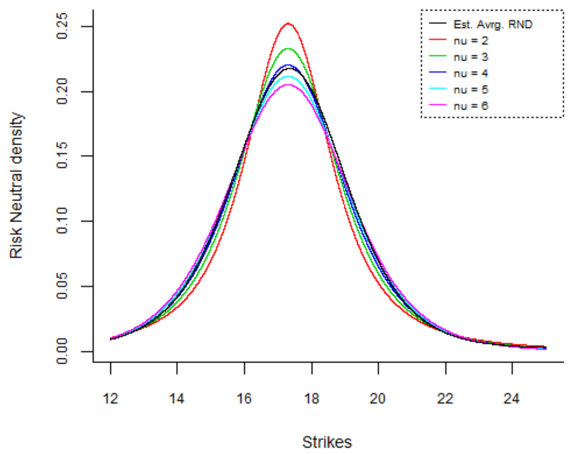
B.1.3/ Student distribution fitting

Degrees of freedom	*σ (visual approximation)*	*σ (optimized fit)*
2	$1.625	$1.4032
3	$1.700	$1.5794
4	$1.720	$1.7043
5	$1.750	$1.7963
6	$1.765	$1.8668

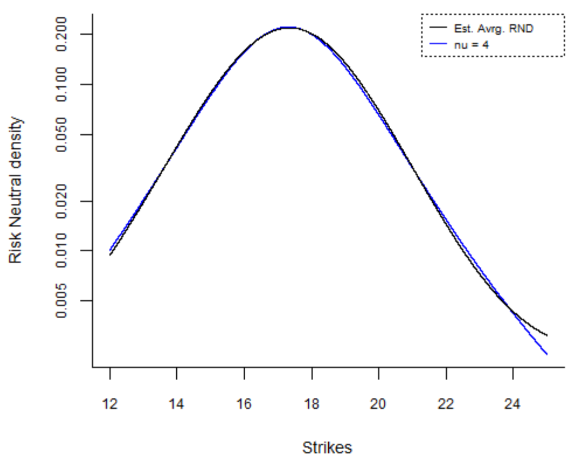
Table A.19: *Optimal scale parameters for fitted non-standardized Student distributions*

Degrees of freedom	Sum of squared errors	Sum of log squared errors
2	24.64	4,430.05
3	5.93	928.43
4	0.98	439.79
5	1.23	1,003.90
6	3.40	1,922.32

Table A.20: *Errors for each fitted non-standardized Student distribution*



Graph A.9*: Comparison of average Risk Neutral Density and fitted non-standardized Student distribution*



Graph A.10*: Comparison of average Risk Neutral Density and best fitted non-standardized Student distribution in terms on log SSE, logarithmic scale*

Option pricing

We explained in section III.4.B of this paper that the distribution of dividends by the underlying stock gives rise to a price difference between a European call option and its American counterpart. We also commented that there are some conditions or computational tricks that allow us to compute an American call option price using techniques for evaluating European call options.

In the case of Citigroup, we showed that the following holds: $D \leq k\left(1 - e^{-r(T-t_D)}\right)$, where t_D is the payment date of the dividend. In such setting, we stated that the price of an American call is equal to the price of its European counterpart. However, this upper bound on the dividend amount does not hold for Bank of America Merrill Lynch. Nonetheless, the dividend to be paid on date t_D – June 21st, 2014 – is equal to \$0.01, which corresponds to 0.058% of the stock price on April 1st and 0.10% of the lowest strike – \$10. Therefore we believe that the theoretical price difference between the American and the European call arising from this dividend payment should be negligible; we think that the European call price that we can estimate with our RN distribution is a satisfactorily good approximation of the American call price so as to assess the accuracy of our calibrated Student distribution.
We have tested this assertion by computing prices for the following options using the `` `RQuantLib` `` package:

- The price of a European call written on Bank of America Merrill Lynch using the Black-Scholes formula, with strike equal to the current price – \$46.55. We linearly interpolated the implied volatilities for strikes \$46 and \$47 from OptionMetrics data. All other parameters have remained unchanged, particularly the annual dividend yield for Bank of America, 0.869%.
- The price of an American call written on Bank of America Merrill Lynch using the Barone-Adesi-Whaley method, with all parameters set as for the European option above.

We found a price difference of just 0.15% between the two prices.

We repeated the same test but this time using the lowest strike for which there still is an implied volatility figure supplied by OptionMetrics – \$12. We found a price difference of 0.22%. Finally, we did the same test again but for the highest possible strike – \$22 – and the price difference was 0.38%.

These results seem to confirm our intuition that, considering the low amounts paid out as dividends by Bank of America Merrill Lynch, American call prices can be reasonably approached by their European counterparts' prices.

We show our call pricing results hereafter. We only show estimated prices through our RN distributions as a percentage over the true market price, as well as relevant statistics.

Strike k	**Estimated $\hat{c}_k$**				
	η=2	**η=3**	**η=4**	**η=5**	**η=6**
\$12	1.52%	**-0.56%**	-1.06%	-1.26%	-1.36%
\$13	2.75%	**-0.03%**	-0.73%	-1.02%	-1.17%
\$14	4.50%	0.66%	**-0.32%**	-0.72%	-0.90%
\$15	5.85%	**0.47%**	-0.84%	-1.29%	-1.45%
\$16	7.92%	**0.19%**	-1.36%	-1.69%	-1.65%
\$17	10.66%	**-1.02%**	-2.86%	-2.88%	-2.42%
\$18	22.40%	**0.61%**	-3.11%	-3.41%	-2.78%
\$19	60.46%	13.23%	2.81%	**-0.04%**	-0.53%
\$20	141.26%	41.19%	16.07%	6.81%	**2.89%**
\$21	283.69%	86.46%	35.85%	15.38%	**5.23%**
\$22	703.60%	229.20%	111.20%	63.20%	**38.80%**

Table A.21: *Estimated call prices as a percentage of market price for each fitted non-standardized Student distribution*

Statistic	η=2	η=3	η=4	η=5	η=6
Average	113.15%	33.97%	16.02%	8.88%	5.38%
St. Dev.	2.14	0.70	0.33	0.19	0.11
Maximum	703.60%	229.20%	111.20%	63.20%	38.80%
Minimum	1.52%	0.03%	0.32%	0.04%	0.53%

Table A.22: *Relative error statistics for estimated call prices for each fitted non-standardized Student distribution*

Statistic	η=2	η=3	η=4	η=5	η=6
Average	113.15%	33.67%	14.15%	6.65%	3.15%
St. Dev.	2.14	0.70	0.34	0.20	0.12
Maximum	703.60%	229.20%	111.20%	63.20%	38.80%
Minimum	1.52%	-1.02%	-3.11%	-3.41%	-2.78%
Best dist.	0	6	1	1	3

Table A.23: *Absolute error statistics for estimated call prices for each fitted non-standardized Student distribution*

Moreover, following the same argument as for Citigroup, we believe that the distortion in an American put price arising from the payment of the 1 cent dividend should be negligible, so we have undertaken pricing computations for put prices also. Results are showed hereafter.

Strike k	**Estimated $\hat{p}_k$**				
	η=2	**η=3**	**η=4**	**η=5**	**η=6**
$13	302.00%	119.50%	56.00%	27.25%	**12.00%**
$14	156.47%	60.12%	26.94%	12.47%	**5.29%**
$15	77.83%	28.29%	12.63%	6.80%	**4.63%**
$16	34.54%	11.38%	5.46%	**4.20%**	4.48%
$17	12.54%	2.23%	**0.28%**	0.43%	1.13%
$18	8.51%	2.36%	1.08%	**1.03%**	1.34%
$19	7.60%	3.22%	1.98%	1.62%	**1.57%**
$20	5.92%	2.86%	1.85%	1.44%	**1.26%**
$21	3.56%	1.46%	0.73%	0.41%	**0.24%**
$22	2.86%	1.39%	0.88%	0.65%	**0.53%**
$23	1.76%	0.72%	0.37%	0.22%	**0.14%**
$24	1.13%	0.38%	0.14%	0.04%	**-0.02%**
$25	0.74%	0.19%	**0.02%**	-0.05%	-0.08%

Table A.24: *Estimated put prices as a percentage of market price for each fitted non-standardized Student distribution*

Statistic	η=2	η=3	η=4	η=5	η=6
Average	55.78%	21.23%	9.84%	5.14%	2.96%
St. Dev.	0.94	0.37	0.17	0.08	0.04
Maximum	302.00%	119.50%	56.00%	27.25%	12.00%
Minimum	1.76%	0.72%	0.28%	0.22%	0.14%

Table A.25: *Relative error statistics for estimated put prices for each fitted non-standardized Student distribution*

Statistic	η=2	η=3	η=4	η=5	η=6
Average	55.78%	21.23%	9.84%	5.14%	2.96%
St. Dev.	0.94	0.37	0.17	0.08	0.04
Maximum	302.00%	119.50%	56.00%	27.25%	12.00%
Minimum	1.76%	0.72%	0.28%	0.22%	0.14%
Best dist.	0	0	2	2	9

Table A.26: *Absolute error statistics for estimated put prices for each fitted non-standardized Student distribution*

Recall that by no arbitrage, the following condition holds – see main appendix with Citigroup's put pricing results: $k_2 > k_1 \Rightarrow p_{k_2} > p_{k_1}$.
Equivalently:

$$k_2 < k_1 \Rightarrow p_{k_2} < p_{k_1}$$

$$k_2 < k_1 \Rightarrow c_{k_2} > c_{k_1}$$

$$k_2 > k_1 \Rightarrow c_{k_2} < c_{k_1}.$$

We have estimated out-of-sample option prices by ignoring some options which market prices do not comply with these conditions – \$25 and \$26 calls, \$10 and \$11 puts.

Strike k	Estimated $\hat{c}_k$				
	η=2	η=3	η=4	η=5	η=6
\$10	1.14%	**-0.15%**	-0.41%	-0.51%	-0.56%
\$11	1.65%	**0.03%**	-0.34%	-0.48%	-0.55%
\$23	1018.00%	293.33%	122.67%	56.00%	**22.67%**
\$24	2776.00%	782.00%	344.00%	182.00%	**106.00%**

Table A.27: *Estimated call prices for testing options as a percentage of market price excluding arbitrageable prices*

Strike k	Estimated $\hat{p}_k$				
	η=2	η=3	η=4	η=5	η=6
\$11	538.00%	180.00%	62.00%	**10.67%**	-16.00%
\$12	513.50%	198.00%	90.50%	42.00%	**16.50%**
\$26	0.78%	0.37%	0.25%	0.21%	**0.18%**

Table A.28: *Estimated put prices for testing options as a percentage of market price excluding arbitrageable prices*

Strike k	Estimated option prices				
	η=2	η=3	η=4	η=5	η=6
\$17 call	10.66%	**-1.02%**	-2.86%	-2.88%	-2.42%
\$18 call	22.40%	**0.61%**	-3.11%	-3.41%	-2.78%
\$17 put	12.54%	2.23%	**0.28%**	0.43%	1.13%
\$18 put	8.51%	2.36%	1.08%	**1.03%**	1.34%
Average	13.53%	1.05%	-1.15%	-1.21%	**-0.68%**
Abs. average	13.53%	**1.56%**	1.83%	1.94%	1.92%

Table A.29: *Estimated option prices for ATM options as a percentage of market price*

Strike k	Estimated option prices				
	η=2	η=3	η=4	η=5	η=6
$18 call	22.40%	**0.61%**	-3.11%	-3.41%	-2.78%
$19 call	60.46%	13.23%	2.81%	**-0.04%**	-0.53%
$20 call	141.26%	41.19%	16.07%	6.81%	**2.89%**
$21 call	283.69%	86.46%	35.85%	15.38%	**5.23%**
$22 call	703.60%	229.20%	111.20%	63.20%	**38.80%**
$23 call	1018.00%	293.33%	122.67%	56.00%	**22.67%**
$24 call	2776.00%	782.00%	344.00%	182.00%	**106.00%**
Average	715.06%	206.57%	89.93%	45.71%	**24.61%**
Abs. average	715.06%	206.57%	90.82%	46.69%	**25.56%**
Best dist.	0	1	0	1	**5**

Table A.30: *Estimated option prices for OTM call options as a percentage of market price*

Risk neutral default probabilities

Degrees of freedom	$\mathbb{P}(X \leq 0)$
2	0.3255%
3	0.0814%
4	0.0265%
5	0.0102%
6	0.0045%

Table A.31: *Three-month risk neutral probabilities of default for fitted non-standardized Student distributions*

Degrees of freedom	Annualized probabilities	Closest S&P rating
2	1.2957%	BB- (1.09%)
3	0.3252%	BBB- (0.30%)
4	0.1060%	BBB+ (0.13%)
5	0.0408%	AA- (0.03%)
6	0.0180%	AA (0.02%)

Table A.32: *One-year risk neutral probabilities of default for fitted non-standardized Student distributions assuming independence and closest matching S&P rating*

Degrees of freedom	Annualized probabilities	Closest S&P rating
2	1.0644%	BB- (1.09%)
3	0.4655%	BB+ (0.40%)
4	0.2574%	BBB- (0.30%)
5	0.1638%	BBB (0.19%)
6	0.1145%	BBB+ (0.13%)

Table A.33: *One-year risk neutral probabilities of default for fitted non-standardized Student distributions assuming non independence and closest matching S&P rating*

In 2014, S&P had a LT local issuer rating of “A-” for Bank of America Corp. The Global Corporate Average Cumulative Default Rate for a one-year horizon corresponding to an “A-” rating is 0.08% – the probability has been estimated by S&P using data from 1981 to 2014.

Degrees of freedom	Annualized probabilities relative to S&P's one-year default probability
2	16.20
3	4.07
4	1.33
5	0.51
6	0.23

Table A.34: *Comparison of fitted default rates (independence hypothesis) and S&P default rates*

Degrees of freedom	Annualized probabilities relative to S&P's one-year default probability
2	13.31
3	5.82
4	3.22
5	2.05
6	1.43

Table A.35: *Comparison of fitted default rates (non independence hypothesis) and S&P default rates*

Risk neutral quantiles

Time to maturity for our panel of options is equal to 0.299, but based on a full year – i.e. taking into account weekend days. If we make the standard assumption that a year has 252 working days, multiplying this figure by 0.299 we obtain a time to maturity of 75 business days. Therefore to compute Bank of America Merrill Lynch's quarterly stock returns we consider a 74-day lag:

$$R_t^{Trimestral} = \frac{S_t - S_{t-74}}{S_{t-74}}.$$

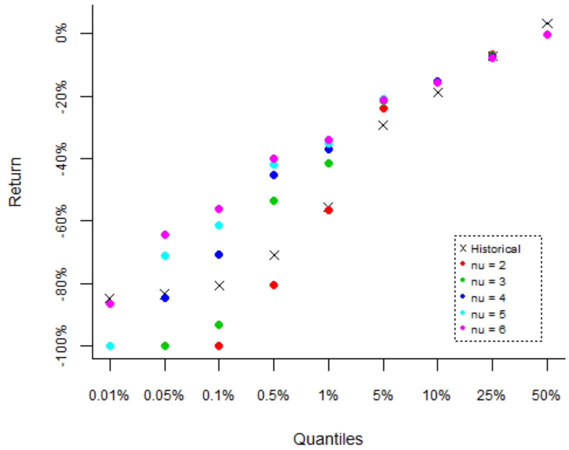
We have computed quantiles for the following stock price quarterly returns time series, ending on April 1st, 2014: 25 years (6,301 observations).

Quantile	25yr	η=2	η=3	η=4	η=5	η=6
0.01%	*-84.77%*	-100%	-100%	-100%	-100%	**-86.59%**
0.05%	*-83.52%*	-100%	-100%	**-84.82%**	-71.35%	-64.34%
0.1%	*-80.67%*	-100%	-93.23%	**-70.69%**	-61.24%	-56.26%
0.5%	*-71.09%*	**-80.51%**	-53.39%	-45.44%	-41.96%	-40.11%
1%	*-55.64%*	**-56.55%**	-41.55%	-37.02%	-35.05%	-34.03%
5%	*-29.35%*	**-23.82%**	-21.63%	-21.14%	-21.07%	-21.11%
10%	*-18.87%*	-15.45%	-15.11%	-15.26%	-15.48%	**-15.69%**
25%	*-7.32%*	-6.80%	-7.16%	**-7.47%**	-7.72%	-7.92%
50%	*3.33%*	-0.19%	-0.19%	-0.19%	-0.19%	-0.19%

Table A.36: *Relevant quantiles for fitted non-standardized Student distributions*
Together with historical quantiles for 25-year returns

Degrees of freedom	Sum of squared errors
2	0.1022
3	0.1259
4	0.1431
5	0.2123
6	0.2485

Table A.37: *Sum of squared errors for each fitted non-standardized Student distribution*



Graphic A.11*: Historical and distributional quantiles for selected distributions*

As for Citigroup's, we observe the same phenomenon happening: except for the Student distribution with *η=2*, RN quantiles for negative returns are generally above historical quantiles. The historical distribution seems to assign a greater probability mass to extreme downward events than our RN distributions, whereas we were expecting the reverse to happen. This result would tend to confirm the idea that market participants underestimate the likelihood of crashes and crisis.

B.2/ Goldman Sachs

In this section, we have undertaken the same procedures as previously but applied to options written on Goldman Sachs stock. Data is from April 1st, 2014. Goldman Sachs shares closed at $165.92 on that day.

B.2.1/ Data

Selected data for Goldman Sachs has better characteristics than that of Citigroup or Bank of America Merrill Lynch: while the average bid-offer spread for all options retained for the estimation procedure were respectively 8.67% and 7.64%, the average for Goldman Sachs is just 5.12%, which should indicate that Goldman Sachs options are more liquid and their prices are better gauges of market participants' expectations.

Calls				Puts			
Strike	*Bid-offer spread*	*Mid*	*Implied volatility*	*Strike*	*Bid-offer spread*	*Mid*	*Implied volatility*
85	4.9%	80.78	-	85	100.0%	0.03	45.25%
90	5.1%	75.85	-	90	300.0%	0.075	46.05%
95	4.9%	71.05	47.08%	95	300.0%	0.075	42.29%
100	5.9%	65.88	-	100	100.0%	0.075	38.71%
105	6.4%	60.88	-	105	112.5%	0.125	37.71%
110	**6.1%**	**56.10**	**37.58%**	110	162.5%	0.145	34.97%
115	**6.9%**	**51.15**	**35.45%**	115	141.7%	0.205	33.27%
120	**1.1%**	**46.30**	**34.90%**	120	94.4%	0.265	31.14%
125	**6.9%**	**41.08**	**26.50%**	125	60.7%	0.365	29.37%
130	**1.4%**	**36.45**	**29.42%**	130	40.5%	0.505	27.64%
135	**1.4%**	**31.63**	**27.43%**	**135**	**8.1%**	**0.645**	**25.43%**
140	**1.7%**	**26.88**	**25.53%**	**140**	**6.6%**	**0.94**	**24.00%**
145	**2.3%**	**22.30**	**23.97%**	**145**	**4.4%**	**1.4**	**22.75%**
150	**2.5%**	**17.93**	**22.54%**	**150**	**1.4%**	**2.115**	**21.72%**
155	**1.4%**	**13.90**	**21.37%**	**155**	**4.9%**	**3.125**	**20.67%**

160	**1.0%**	**10.35**	**20.49%**
165	**2.1%**	**7.33**	**19.69%**
170	**3.1%**	**4.93**	**19.08%**
175	**4.9%**	**3.13**	**18.57%**
180	**3.7%**	**1.91**	**18.30%**
185	**6.5%**	**1.12**	**18.14%**
190	**11.3%**	**0.66**	**18.24%**
195	36.4%	0.39	18.50%
200	83.3%	0.26	19.12%
205	53.3%	0.19	20.09%
210	240.0%	0.11	20.22%
215	140.0%	0.09	21.15%
220	900.0%	0.06	21.53%
225	-	0.05	22.46%
230	-	0.04	23.21%
235	-	0.03	24.14%
240	-	0.03	24.94%
245	-	0.03	26.16%

160	**4.4%**	**4.6**	**19.84%**
165	**2.3%**	**6.625**	**19.14%**
170	**1.6%**	**9.225**	**18.48%**
175	**2.0%**	**12.475**	**18.03%**
180	**1.6%**	**16.225**	**17.49%**
185	**14.4%**	**21.175**	**20.60%**
190	**12.8%**	**25.325**	**19.13%**
195	**12.1%**	**30.275**	**21.35%**
200	1.3%	34.525	-
205	**9.2%**	**40**	**23.58%**
210	**8.2%**	**44.925**	**24.79%**
215	**7.0%**	**49.95**	**27.04%**
220	**6.5%**	**54.925**	**28.61%**
225	**5.7%**	**59.95**	**30.76%**
230	**6.2%**	**64.65**	**26.57%**
235	**5.8%**	**69.675**	**28.89%**
240	4.5%	74.425	-
245	**5.1%**	**79.675**	**31.86%**

***Tables A.38 and A.39**: Option prices data for Goldman Sachs on April 1st, for a 109 days maturity (2014)*

B.2.2/ Risk neutral density estimation

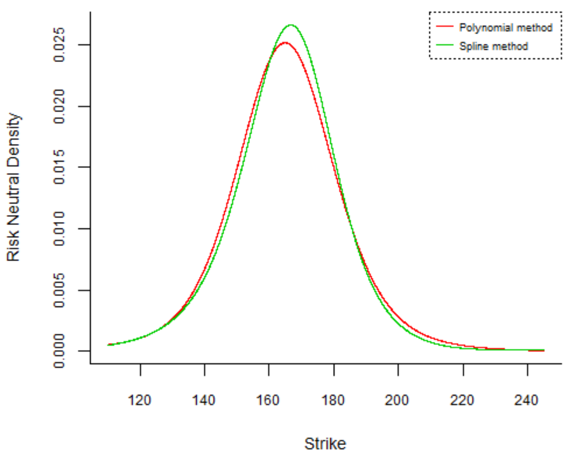
As for Bank of America Merrill Lynch, polynomial interpolation is undertaken using a 3 degree polynomial. Moreover, we have ignored results obtained through the Nadaraya-Watson method as we obtained a RN density that was negative on some subsets of prices.

Method	Polynomial interpolation	Spline smoothing	Nadaraya-Watson estimator: h = h*	Nadaraya-Watson estimator: h = 30
SSE for implied volatility estimation	0.0064770	0.0046727	-	-

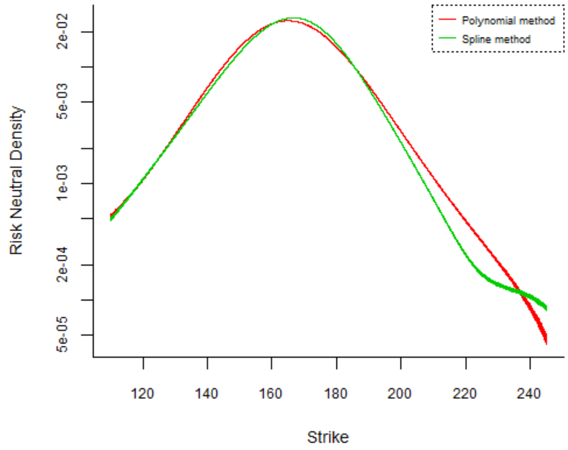
***Table A.40**: Sum of Squared Errors for all implied volatility estimation methods*

Method	Polynomial interpolation	Spline smoothing	Nadaraya-Watson estimator: h = h*	Nadaraya-Watson estimator: h = 30
Density integral estimate	99.61%	98.70%	-	-

***Table A.41**: Density integral estimates for polynomial, spline and Nadaraya-Watson methods*



***Graph A.12**: Estimated RNDs for Goldman Sachs by polynomial and spline interpolation*



***Graph A.13**: Estimated RNDs for Goldman Sachs by polynomial and spline interpolation, logarithmic scale*

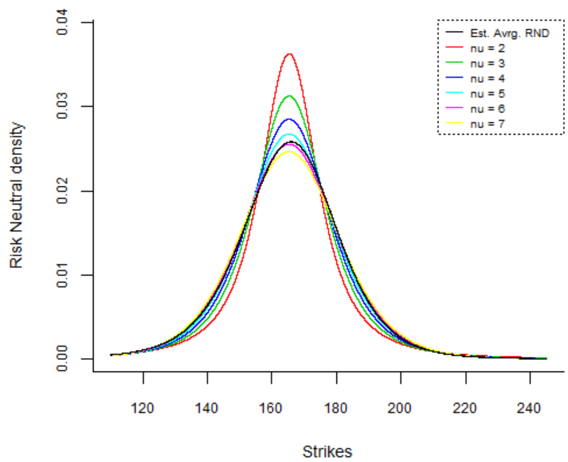
B.2.3/ Student distribution fitting

Degrees of freedom	*σ (visual approximation)*	*σ (optimized fit)*
2	$13.50	$9.7392
3	$10.00	$11.7447
4	$12.50	$13.1450
5	$14.50	$14.1785
6	$15.00	$14.9740
7	$15.20	$15.6058

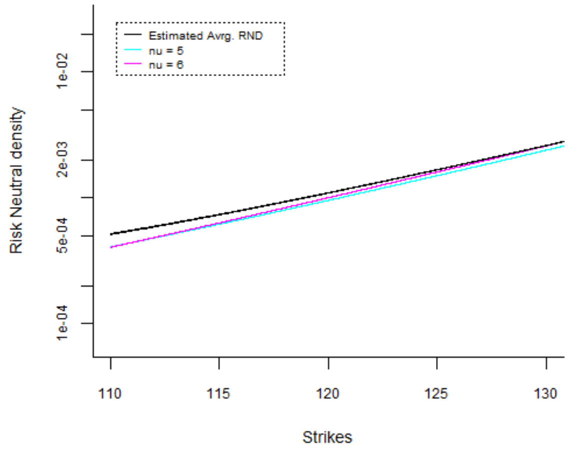
***Table A.42:** Optimal scale parameters for fitted non-standardized Student distributions*

Degrees of freedom	Sum of squared errors	Sum of log squared errors
2	12.57	223,219.00
3	3.98	80,919.78
4	1.07	26,571.31
5	0.18	9,865.59
6	0.07	10,320.62
7	0.30	18,748.43

***Table A.43:** Errors for each fitted non-standardized Student distribution*



***Graph A.14**: Comparison of average Risk Neutral Density and fitted non-standardized Student distribution*



***Graph A.15**: Comparison of average Risk Neutral Density and best fitted non-standardized Student distributions in terms on SSE (in terms of squared errors and squared log-errors), logarithmic scale on the range $110-$130*

Option pricing

As for Bank of America, the dividend payment made by Goldman Sachs on May 28th does not respect the upper bound constraint that would allow us to price American options as European ones.

In the case of Goldman Sachs, the dividend payment is equal to $0.55, which represents 0.331% of the stock price on April 1st – against 0.058% in the case of Bank of America – and 0.647% of the lowest strike, $85 – against 0.100% for Bank of America.

Although these values increase with respect to Bank of America Merrill Lynch, they still remain relatively low. We have decided to retain again the option pricing test to assess the accuracy of our procedure. Results are shown hereafter.

Strike k	Estimated $\hat{c}_k$					
	η=2	η=3	η=4	η=5	η=6	η=7
$110	**0.02%**	-1.00%	-1.22%	-1.30%	-1.35%	-1.37%
$115	**0.09%**	-1.08%	-1.35%	-1.46%	-1.51%	-1.53%
$120	**0.00%**	-1.37%	-1.69%	-1.82%	-1.88%	-1.91%
$125	0.85%	**-0.77%**	-1.16%	-1.31%	-1.38%	-1.42%
$130	**0.35%**	-1.55%	-2.01%	-2.19%	-2.27%	-2.30%
$135	**0.48%**	-1.79%	-2.33%	-2.52%	-2.59%	-2.61%
$140	**0.62%**	-2.09%	-2.70%	-2.88%	-2.91%	-2.89%
$145	**0.47%**	-2.75%	-3.37%	-3.45%	-3.38%	-3.25%
$150	**0.06%**	-3.70%	-4.19%	-4.05%	-3.76%	-3.43%
$155	**-1.12%**	-5.32%	-5.42%	-4.84%	-4.15%	-3.49%
$160	**-3.14%**	-7.69%	-7.09%	-5.82%	-4.53%	-3.37%
$165	-3.96%	-9.53%	-8.21%	-6.09%	-4.05%	**-2.25%**
$170	**-0.03%**	-9.34%	-7.94%	-5.16%	-2.39%	0.10%
$175	14.22%	-4.03%	-4.15%	**-1.35%**	1.88%	4.96%
$180	42.70%	7.77%	**3.61%**	5.24%	8.28%	11.55%
$185	94.16%	29.98%	18.06%	**16.72%**	18.54%	21.37%
$190	172.76%	61.54%	36.85%	30.09%	**29.18%**	30.50%

***Table A.44:** Estimated call prices as a percentage of market price for each fitted non-standardized Student distribution*

Statistic	η=2	η=3	η=4	η=5	η=6	η=7
Average	19.71%	8.90%	6.55%	5.66%	5.53%	5.78%
St. Dev.	0.45	0.15	0.09	0.07	0.07	0.08
Maximum	172.76%	61.54%	36.85%	30.09%	29.18%	30.50%
Minimum	0.00%	0.77%	1.16%	1.30%	1.35%	0.10%

Table A.45: *Relative error statistics for estimated call prices for each fitted non-standardized Student distribution*

Statistic	η=2	η=3	η=4	η=5	η=6	η=7
Average	18.74%	2.78%	0.33%	0.46%	1.28%	2.27%
St. Dev.	0.47	0.18	0.11	0.09	0.09	0.10
Maximum	172.76%	61.54%	36.85%	30.09%	29.18%	30.50%
Minimum	-3.96%	-9.53%	-8.21%	-6.09%	-4.53%	-3.49%
Best dist.	11	1	1	2	1	1

Table A.46: *Absolute error statistics for estimated call prices for each fitted non-standardized Student distribution*

Again, in order to test whether our technique yields prices very far away from market prices, we have also computed put prices despite the presence of dividend payouts – given the low amount of the dividend, distortions should be small. Results are showed hereafter.

Strike k	Estimated $\hat{p}_k$					
	η=2	η=3	η=4	η=5	η=6	η=7
\$135	86.68%	14.64%	**-8.03%**	-16.62%	-19.94%	-20.93%
\$140	55.86%	**5.10%**	-9.60%	-14.07%	-14.85%	-14.13%
\$145	30.52%	**-2.91%**	-10.91%	-11.96%	-10.71%	-8.67%
\$150	11.30%	-8.71%	-11.62%	-10.24%	-7.69%	**-4.92%**
\$155	1.08%	-9.58%	-9.19%	-6.47%	-3.37%	**-0.41%**
\$160	-3.55%	-8.33%	-6.44%	-3.47%	**-0.55%**	2.07%
\$165	-2.75%	-5.13%	-3.27%	**-0.86%**	1.41%	3.41%
\$170	1.11%	-1.13%	**-0.11%**	1.43%	2.92%	4.25%
\$175	3.97%	**1.41%**	1.59%	2.33%	3.15%	3.92%
\$180	5.49%	2.94%	**2.61%**	2.83%	3.19%	3.58%
\$185	1.81%	**-0.38%**	-0.89%	-0.94%	-0.83%	-0.68%
\$190	3.37%	1.48%	0.94%	0.79%	**0.77%**	0.80%
\$195	2.07%	0.53%	**0.05%**	-0.13%	-0.20%	-0.21%
\$205	1.34%	0.31%	**-0.03%**	-0.18%	-0.25%	-0.29%
\$210	1.08%	0.23%	**-0.05%**	-0.17%	-0.24%	-0.27%
\$215	0.71%	**0.01%**	-0.22%	-0.32%	-0.37%	-0.40%
\$220	0.54%	**-0.05%**	-0.24%	-0.32%	-0.36%	-0.38%
\$225	0.33%	**-0.16%**	-0.32%	-0.38%	-0.41%	-0.43%
\$230	0.67%	0.26%	0.13%	0.08%	0.06%	**0.04%**
\$235	0.51%	0.16%	0.06%	0.02%	**0.00%**	-0.02%
\$245	0.33%	0.08%	**0.01%**	-0.02%	-0.03%	-0.03%

Table A.47: *Estimated put prices as a percentage of market price for each fitted non-standardized Student distribution*

In this case we observe that the best performing distribution varies wildly from strike to strike, particularly for the range $135-$190. This phenomenon might be linked to the fact that we are pricing American puts as if they were European; it could also mean that the empirical RN distribution is particularly irregular.

Statistic	η=2	η=3	η=4	η=5	η=6	η=7
Average	10.24%	3.02%	3.16%	3.50%	3.39%	3.33%
St. Dev.	0.21	0.04	0.04	0.05	0.05	0.05
Maximum	86.68%	14.64%	11.62%	16.62%	19.94%	20.93%
Minimum	0.33%	0.01%	0.01%	0.02%	0.00%	0.02%

Table A.48: *Relative error statistics for estimated put prices for each fitted non-standardized Student distribution*

Statistic	η=2	η=3	η=4	η=5	η=6	η=7
Average	9.64%	-0.44%	-2.64%	-2.79%	-2.30%	-1.60%
St. Dev.	0.22	0.05	0.04	0.05	0.06	0.06
Maximum	86.68%	14.64%	2.61%	2.83%	3.19%	4.25%
Minimum	-3.55%	-9.58%	-11.62%	-16.62%	-19.94%	-20.93%
Best dist.	0	7	7	1	3	3

Table A.49: *Absolute error statistics for estimated put prices for each fitted non-standardized Student distribution*

We have estimated out-of-sample option prices by ignoring some options which market prices do not comply with arbitrage conditions – $235, $240 and $245 calls, $85, $90, $95 and $100 puts.

Strike k	Estimated $\hat{c}_k$					
	η=2	η=3	η=4	η=5	η=6	η=7
$85	**0.07%**	-0.49%	-0.59%	-0.62%	-0.63%	-0.64%
$90	**0.04%**	-0.58%	-0.70%	-0.74%	-0.75%	-0.76%
$95	**-0.17%**	-0.87%	-1.00%	-1.05%	-1.07%	-1.08%
$100	**0.15%**	-0.63%	-0.79%	-0.85%	-0.88%	-0.89%
$105	**0.27%**	-0.63%	-0.82%	-0.89%	-0.92%	-0.94%
$195	289.10%	104.51%	60.46%	45.31%	39.85%	**38.41%**
$200	405.85%	137.81%	72.38%	47.77%	36.92%	**31.92%**
$205	510.11%	158.53%	73.11%	40.00%	24.26%	**15.95%**
$210	841.36%	262.18%	124.55%	71.00%	44.91%	**30.45%**
$215	939.00%	265.44%	110.33%	50.78%	21.78%	**5.44%**
$220	1320.50%	359.67%	146.17%	66.33%	27.83%	**6.17%**
$225	1465.40%	368.60%	134.60%	49.40%	**9.20%**	-13.00%
$230	1708.75%	403.50%	136.00%	42.25%	**-1.00%**	-24.50%

Table A.50: *Estimated call prices for testing options as a percentage of market price excluding arbitrageable prices*

Strike k	Estimated $\hat{p}_k$					
	η=2	η=3	η=4	η=5	η=6	η=7
$105	291.76%	58.00%	**-13.92%**	-43.68%	-58.48%	-66.88%
$110	284.69%	64.00%	**-5.52%**	-34.83%	-49.72%	-58.28%
$115	211.61%	41.12%	**-13.71%**	-37.17%	-49.17%	-56.10%
$120	178.00%	34.49%	**-12.34%**	-32.57%	-42.87%	-48.68%
$125	134.82%	22.14%	**-14.88%**	-30.74%	-38.60%	-42.85%

$130	99.70%	**12.46%**	-15.98%	-27.72%	-33.17%	-35.78%
$200	3.40%	2.11%	1.69%	1.52%	1.44%	**1.40%**
$240	0.75%	0.45%	0.37%	0.33%	0.32%	**0.31%**

Table A.51: *estimated put prices for testing options as a percentage of market price excluding arbitrageable prices*

Strike k	**Estimated option prices**					
	η=2	**η=3**	**η=4**	**η=5**	**η=6**	**η=7**
$165 call	-3.96%	-9.53%	-8.21%	-6.09%	-4.05%	**-2.25%**
$170 call	**-0.03%**	-9.34%	-7.94%	-5.16%	-2.39%	0.10%
$165 put	-2.75%	-5.13%	-3.27%	**-0.86%**	1.41%	3.41%
$170 put	1.11%	-1.13%	**-0.11%**	1.43%	2.92%	4.25%
Average	-1.41%	-6.28%	-4.88%	-2.67%	**-0.53%**	1.38%
Abs. avrg.	**1.96%**	6.28%	4.88%	3.39%	2.69%	2.50%

Table A.52: *Estimated option prices for ATM options as a percentage of market price*

Strike k	**Estimated option prices**					
	η=2	**η=3**	**η=4**	**η=5**	**η=6**	**η=7**
$170 call	**-0.03%**	-9.34%	-7.94%	-5.16%	-2.39%	0.10%
$175 call	14.22%	-4.03%	-4.15%	**-1.35%**	1.88%	4.96%
$180 call	42.70%	7.77%	**3.61%**	5.24%	8.28%	11.55%
$185 call	94.16%	29.98%	18.06%	**16.72%**	18.54%	21.37%
$190 call	172.76%	61.54%	36.85%	30.09%	**29.18%**	30.50%
$195 call	289.10%	104.51%	60.46%	45.31%	39.85%	**38.41%**
$200 call	405.85%	137.81%	72.38%	47.77%	36.92%	**31.92%**
$205 call	510.11%	158.53%	73.11%	40.00%	24.26%	**15.95%**
$210 call	841.36%	262.18%	124.55%	71.00%	44.91%	**30.45%**
$215 call	939.00%	265.44%	110.33%	50.78%	21.78%	**5.44%**
$220 call	1320.50%	359.67%	146.17%	66.33%	27.83%	**6.17%**
$225 call	1465.40%	368.60%	134.60%	49.40%	**9.20%**	-13.00%
$230 call	1708.75%	403.50%	136.00%	42.25%	**-1.00%**	-24.50%
Average	600.30%	165.09%	69.54%	35.26%	19.94%	**12.26%**
Abs. avrg.	600.30%	167.15%	71.40%	36.26%	20.46%	**18.02%**
Best dist.	1	0	1	2	3	**6**

Table A.53: *Estimated option prices for OTM call options as a percentage of market price*

Risk neutral default probabilities

Degrees of freedom	$\mathbb{P}(X \leq 0)$
2	0.1726%
3	0.0388%
4	0.0115%
5	0.0041%
6	0.0016%
7	0.0007%

Table A.54: *Three-month risk neutral probabilities of default for fitted non-standardized Student distributions*

Degrees of freedom	Annualized probabilities	Closest S&P rating
2	0.6886%	BB (0.64%)
3	0.1551%	BBB+ (0.13%)
4	0.0460%	A+ (0.06%)
5	0.0164%	AA (0.02%)
6	0.0064%	AA+, AAA (0.00%)
7	0.0028%	AA+, AAA (0.00%)

Table A.55: *One-year risk neutral probabilities of default for fitted non-standardized Student distributions assuming independence and closest matching S&P rating*

Degrees of freedom	Annualized probabilities	Closest S&P rating
2	0.5705%	BB (0.64%)
3	0.2279%	BBB (0.19%)
4	0.1172%	BBB+ (0.13%)
5	0.0702%	A (0.07%)
6	0.0467%	A+ (0.06%)
7	0.0334%	AA- (0.03%)

Table A.56: *One-year risk neutral probabilities of default for fitted non-standardized Student distributions assuming non independence and closest matching S&P rating*

In 2014, S&P had a LT local issuer rating of “A-” for Goldman Sachs Group Inc. The Global Corporate Average Cumulative Default Rate for a one-year horizon corresponding to an “A-” rating is 0.08% – the probability has been estimated by S&P using data from 1981 to 2014.

Degrees of freedom	Annualized probabilities relative to S&P’s one-year default probability
2	8.61
3	1.94
4	0.58
5	0.21
6	0.08
7	0.04

Table A.57: *Comparison of fitted default rates (independence hypothesis) and S&P default rates*

Degrees of freedom	Annualized probabilities relative to S&P’s one-year default probability
2	7.13
3	2.85
4	1.47
5	0.88
6	0.58
7	0.42

Table A.58: *Comparison of fitted default rates (non independence hypothesis) and S&P default rates*

Risk neutral quantiles

Time to maturity for our panel of options is equal to 0.299, but based on a full year – i.e. taking into account weekend days. If we make the standard assumption that a year has 252 working days, multiplying this figure by 0.299 we obtain a time to maturity of 75 business days. Therefore to compute Citigroup’s quarterly stock returns we consider a 74-day lag:

$$R_t^{Trimestral} = \frac{S_t - S_{t-74}}{S_{t-74}}$$

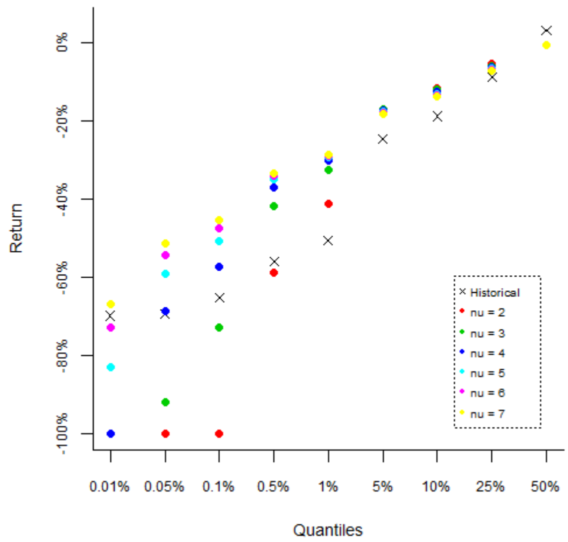
We have computed quantiles for the following stock price quarterly returns time series, ending on April 1st, 2014: approximately 14 years and ¾ (3,679 observations), given that Goldman Sachs stock history began in 1999.

Quantile	14yr ¾	η=2	η=3	η=4	η=5	η=6	η =7
0.01%	*-69.78%*	-100%	-100%	-100%	-83.07%	-72.79%	**-66.80%**
0.05%	*-69.34%*	-100%	-91.85%	**-68.58%**	-59.06%	-54.15%	-51.23%
0.1%	*-65.13%*	-100%	**-72.67%**	-57.20%	-50.73%	-47.37%	-45.38%
0.5%	*-55.94%*	**-58.62%**	-41.71%	-36.84%	-34.82%	-33.83%	-33.28%
1%	*-50.40%*	**-41.25%**	-32.51%	-30.05%	-29.12%	-28.73%	-28.57%
5%	*-24.56%*	-17.51%	-17.03%	-17.26%	-17.59%	-17.90%	**-18.19%**
10%	*-18.67%*	-11.44%	-11.96%	-12.51%	-12.98%	-13.36%	**-13.68%**
25%	*-8.62%*	-5.16%	-5.78%	-6.24%	-6.58%	-6.84%	**-7.06%**
50%	*+3.37%*	-0.37%	-0.37%	-0.37%	-0.37%	-0.37%	-0.37%

Table A.59: *Relevant quantiles for fitted non-standardized Student distributions Together with historical quantiles for 14¾ -year returns*

Degrees of freedom	Sum of squared errors
2	0.3288
3	0.1177
4	0.0811
5	0.0425
6	0.0143
7	0.0049

Table A.60: *Sum of squared errors for each fitted non-standardized Student distribution*



Graph A.16*: Historical and distributional quantiles for selected distributions*

B.3/ Morgan Stanley

In this section, we have undertaken the same procedures as previously but applied to options written on Morgan Stanley stock. Data is from April 1st, 2014. Morgan Stanley shares closed at $31.21 on that day.

B.3.1/ Data

In the case of Morgan Stanley, thanks to the availability of data we have been able to retain only options for which the bid-offer spread was narrower or equal to 35%.

We notice that deep ITM options seem to be highly traded, as the spread diminishes progressively as the strike gets more in the money, both for calls and puts; this was already the case for Bank of America Merrill Lynch. The average bid-offer spread for both ITM calls and puts for Citigroup, Bank of America and Goldman Sachs were respectively 11.01%, 1.38% and 5.03% against 1.92% for Morgan Stanley.

As for the average bid-offer spread for option prices used for the estimation procedure, it is equal to 4.10%, while for Citigroup, Bank of America and Goldman Sachs they were respectively 8.67%, 7.64% and 5.12%.

Calls			
Strike	*Bid-offer spread*	*Mid*	*Implied volatility*
16	1.0%	15.23	-
18	1.1%	13.23	-
19	1.2%	12.23	-
20	1.3%	11.23	-
21	**1.0%**	**10.25**	**38.29%**
22	**1.6%**	**9.28**	**36.94%**
23	**1.2%**	**8.30**	**34.77%**
24	**2.1%**	**7.33**	**32.22%**
25	**2.4%**	**6.38**	**30.38%**
26	**2.8%**	**5.48**	**29.57%**
27	**3.3%**	**4.58**	**27.72%**
28	**4.1%**	**3.78**	**27.22%**
29	**2.3%**	**3.02**	**26.22%**
30	**3.0%**	**2.35**	**25.50%**
31	**3.4%**	**1.77**	**24.91%**
32	**4.0%**	**1.29**	**24.28%**
33	**5.7%**	**0.91**	**23.85%**
34	**6.7%**	**0.62**	**23.56%**
35	**10.3%**	**0.41**	**23.28%**
36	**12.0%**	**0.27**	**23.12%**
37	26.7%	0.17	23.11%
38	40.0%	0.12	23.70%
39	66.7%	0.08	23.99%
40	75.0%	0.06	24.41%
41	150.0%	0.04	24.52%
42	400.0%	0.03	25.71%

Puts			
Strike	*Bid-offer spread*	*Mid*	*Implied volatility*
16	-	0.02	-
18	-	0.025	-
19	400.0%	0.03	42.09%
20	200.0%	0.04	39.83%
21	133.3%	0.05	37.26%
22	50.0%	0.075	35.87%
23	44.4%	0.11	34.46%
24	30.8%	0.15	32.62%
25	**21.1%**	**0.21**	**31.03%**
26	**7.1%**	**0.29**	**29.40%**
27	**2.4%**	**0.415**	**28.19%**
28	**10.3%**	**0.61**	**27.50%**
29	**7.2%**	**0.86**	**26.64%**
30	**5.2%**	**1.19**	**25.88%**
31	**2.5%**	**1.6**	**25.05%**
32	**2.4%**	**2.135**	**24.73%**
33	**1.8%**	**2.755**	**24.32%**
34	**1.4%**	**3.475**	**24.17%**
35	**1.2%**	**4.275**	**24.21%**
36	**3.0%**	**5.075**	**22.82%**
37	**1.7%**	**6**	**23.38%**
38	**1.4%**	**6.95**	**24.01%**
39	**1.9%**	**7.925**	**25.16%**
40	**1.1%**	**8.9**	**25.86%**
41	**1.0%**	**9.9**	**27.89%**
42	**0.9%**	**10.9**	**29.85%**

Tables A.61 and A.62*: Option prices data for Morgan Stanley on April 1st, for a 109 days maturity (2014)*

We had to reprocess the linearly interpolated implied volatility value for $35. Indeed, in the above Table the reader can notice that implied volatility for put options increases at 35$, but then decreases again at $36, while the proper volatility smile starts at $37 – from there on, implied volatility is monotonically increasing. This pattern resulted in a significant jump in our interpolated volatility curve so we decided to take the average of the interpolated values at $34 and $36 and replace the $35 volatility with this value.

B.3.2/ Risk neutral density estimation

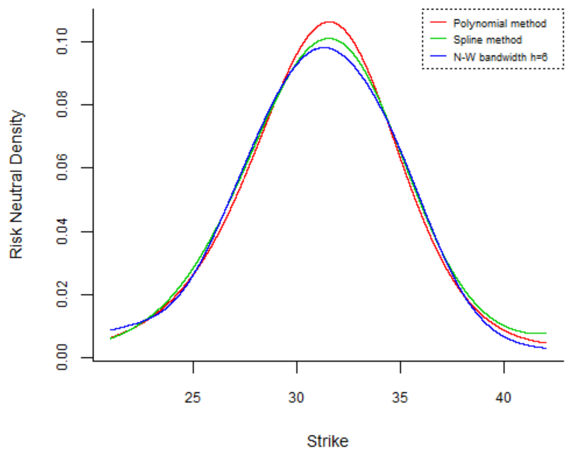
Polynomial interpolation has been undertaken with a 3 degree polynomial – same methodology as for Bank of America and Goldman Sachs. Spline smoothing was implemented with 7 degree polynomials to obtain a smooth enough result.

Method	Polynomial interpolation	Spline smoothing	Nadaraya-Watson estimator: h = h*	Nadaraya-Watson estimator: h = 6
SSE for implied volatility estimation	0.0005346	0.0002398	0.0007010	0.0024369

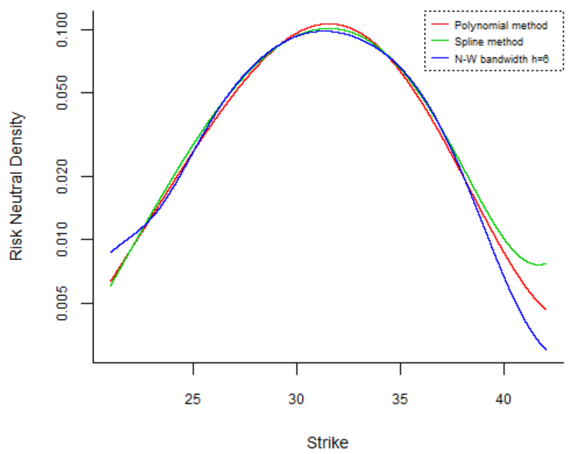
Table A.63*: Sum of Squared Errors for all implied volatility estimation methods*

Method	Polynomial interpolation	Spline smoothing	Nadaraya-Watson estimator: h = h*	Nadaraya-Watson estimator: h = 6
Density integral estimate	97.76%	98.91%	96.04%	96.48%

Table A.64*: Density integral estimates for polynomial, spline and Nadaraya-Watson methods*

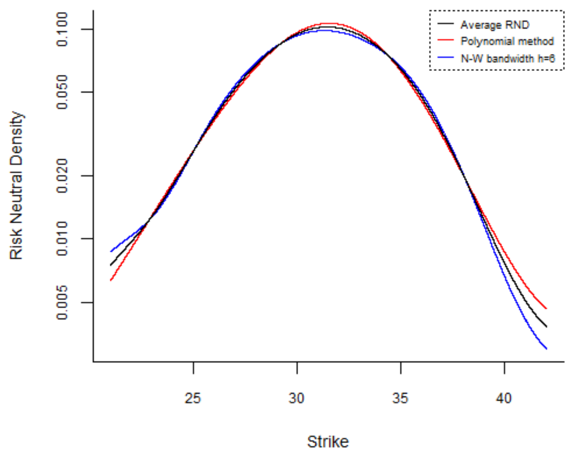


Graphic A.17*: Estimated RNDs for Morgan Stanley*



Graph A.18*: Estimated RNDs for Morgan Stanley, logarithmic scale*

We notice that the Nadaraya-Watson technique seems to yielded smooth enough results for the 1st time, compared to the densities we obtained for Citigroup, Bank of America and Goldman Sachs. On the contrary, the spline method presents an undesirable pattern on the right tail, as the density curve flattens abruptly. To compute the average RN density we have decided to keep densities computed through the polynomial and Nadaraya-Watson methods.



***Graph A.19**: Estimated RNDs and average RND for Morgan Stanley, logarithmic scale*

B.3.3/ Student distribution fitting

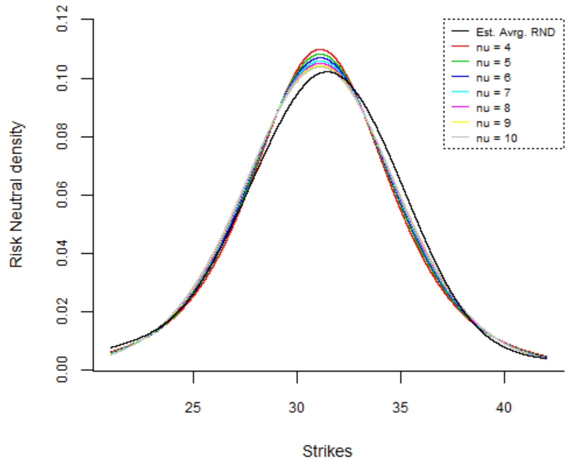
Degrees of freedom	***σ*** *(visual approximation)*	***σ*** *(optimized fit)*
2	$3.25	$3.1194
3	$3.25	$3.2886
4	$3.50	$3.4148
5	$3.50	$3.5084
6	$3.60	$3.5795
7	$3.65	$3.6353
8	$3.675	$3.6799
9	$3.70	$3.7165
10	$3.75	$3.7470

***Table A.65:** Optimal scale parameters for fitted non-standardized Student distributions*

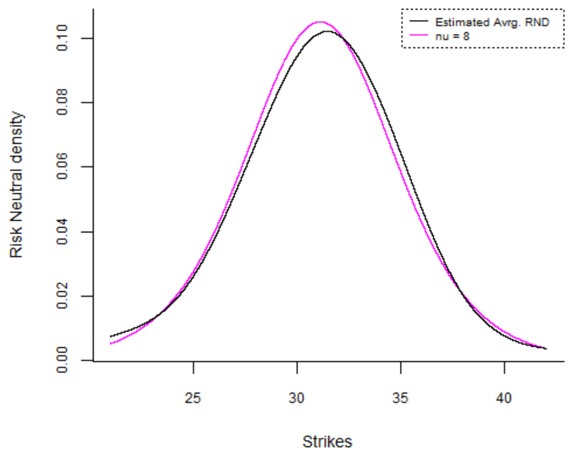
Degrees of freedom	**Sum of squared errors**	**Sum of log squared errors**
2	11.39	7,851.44
3	6.73	4,332.44
4	4.51	2,894.60
5	3.33	2,257.97
6	2.66	1,977.35
7	2.28	1,868.95
8	2.06	1,848.24
9	1.93	1,873.28
10	1.86	1,922.09

***Table A.66:** Errors for each fitted non-standardized Student distribution*

We observe that the sum of squared errors might keep decreasing at η = *11*, however a minimum is reached for log squared errors at 8 degrees of freedom.



***Graph A.20**: Comparison of average Risk Neutral Density and fitted non-standardized Student distribution*



***Graph A.21**: Comparison of average Risk Neutral Density and Best fitted non-standardized Student distribution in terms on SSE*

Option pricing

Morgan Stanley's 10 cent dividend paid on April 28th does not verify the upper bound for the lowest strike used during the estimation procedure – $21. Nevertheless we decide again to undertake the option pricing test: the dividend payment represents only 0.320% of the stock price on April 1st – against 0.058% for Bank of America and 0.331% for Goldman Sachs – and 0.625% of the lowest strike, $16 – against 0.100% for Bank of America and 0.647% for Goldman Sachs.

Strike k	**Estimated $\hat{c}_k$**								
	η=2	**η=3**	**η=4**	**η=5**	**η=6**	**η=7**	**η=8**	**η=9**	**η=10**
21	2.99%	**0.09%**	-0.61%	-0.90%	-1.04%	-1.12%	-1.18%	-1.22%	-1.24%
22	3.54%	**0.17%**	-0.68%	-1.03%	-1.21%	-1.32%	-1.40%	-1.44%	-1.48%
23	4.32%	**0.38%**	-0.66%	-1.10%	-1.33%	-1.47%	-1.56%	-1.63%	-1.68%
24	5.49%	0.81%	**-0.46%**	-1.01%	-1.30%	-1.48%	-1.60%	-1.69%	-1.75%
25	6.86%	1.24%	**-0.32%**	-1.00%	-1.37%	-1.59%	-1.75%	-1.85%	-1.93%
26	8.14%	1.34%	**-0.57%**	-1.40%	-1.86%	-2.14%	-2.32%	-2.45%	-2.54%
27	10.61%	2.25%	**-0.12%**	-1.15%	-1.70%	-2.03%	-2.24%	-2.39%	-2.50%
28	12.46%	2.14%	**-0.76%**	-2.00%	-2.65%	-3.03%	-3.27%	-3.43%	-3.54%
29	15.84%	2.83%	**-0.77%**	-2.28%	-3.04%	-3.47%	-3.73%	-3.91%	-4.02%
30	20.39%	3.66%	**-0.89%**	-2.74%	-3.65%	-4.16%	-4.46%	-4.64%	-4.75%
31	27.46%	5.34%	**-0.62%**	-3.03%	-4.20%	-4.84%	-5.20%	-5.42%	-5.56%

32	40.13%	9.63%	**1.35%**	-2.01%	-3.66%	-4.56%	-5.09%	-5.42%	-5.62%
33	60.33%	16.95%	4.99%	**0.03%**	-2.46%	-3.88%	-4.73%	-5.28%	-5.65%
34	91.74%	28.74%	11.10%	3.58%	**-0.32%**	-2.58%	-4.02%	-4.98%	-5.65%
35	142.24%	48.44%	21.95%	10.44%	4.32%	**0.66%**	-1.73%	-3.37%	-4.54%
36	219.09%	77.85%	38.00%	20.53%	11.09%	5.36%	1.55%	**-1.13%**	-3.06%

Table A.67: *Estimated call prices as a percentage of market price for each fitted non-standardized Student distribution*

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8	η=9	η=10
Average	41.98%	12.62%	5.24%	3.39%	2.83%	2.73%	2.86%	3.14%	3.47%
St. Dev.	0.59	0.21	0.10	0.05	0.02	0.01	0.01	0.02	0.02
Maximum	219.09%	77.85%	38.00%	20.53%	11.09%	5.36%	5.20%	5.42%	5.65%
Minimum	2.99%	0.09%	0.12%	0.03%	0.32%	0.66%	1.18%	1.13%	1.24%

Table A.68: *Relative error statistics for estimated call prices for each fitted non-standardized Student distribution*

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8	η=9	η=10
Average	41.98%	12.62%	4.43%	0.93%	-0.90%	-1.98%	-2.67%	-3.14%	-3.47%
St. Dev.	0.59	0.21	0.11	0.06	0.04	0.02	0.02	0.02	0.02
Maximum	219.09%	77.85%	38.00%	20.53%	11.09%	5.36%	1.55%	-1.13%	-1.24%
Minimum	2.99%	0.09%	-0.89%	-3.03%	-4.20%	-4.84%	-5.20%	-5.42%	-5.65%
Best dist.	0	3	9	1	1	1	0	1	0

Table A.69: *Absolute error statistics for estimated call prices for each fitted non-standardized Student distribution*

Again, in order to test whether our technique yields prices very far away from market prices, we have also computed put prices despite the presence of dividend payouts – given the low amount of the dividend, distortions should be small. Results are showed hereafter.

Strike k	Estimated $\hat{p}_k$								
	η=2	η=3	η=4	η=5	η=6	η=7	η=8	η=9	η=10
25	165.52%	59.10%	19.24%	**0%**	-10.86%	-17.62%	-22.19%	-25.43%	-27.86%
26	129.31%	47.55%	16.83%	**2.03%**	-6.24%	-11.38%	-14.83%	-17.24%	-19.03%
27	93.83%	34.07%	11.73%	**1.13%**	-4.75%	-8.34%	-10.70%	-12.31%	-13.52%
28	62.07%	20.26%	4.87%	**-2.30%**	-6.18%	-8.48%	-9.97%	-10.95%	-11.66%
29	43.63%	13.65%	2.84%	**-2.07%**	-4.65%	-6.13%	-7.05%	-7.65%	-8.05%
30	31.50%	9.87%	2.21%	**-1.18%**	-2.92%	-3.89%	-4.47%	-4.82%	-5.06%
31	24.83%	8.78%	3.15%	0.68%	**-0.57%**	-1.25%	-1.65%	-1.89%	-2.04%
32	19.03%	6.98%	2.73%	0.84%	**-0.11%**	-0.64%	-0.96%	-1.15%	-1.27%
33	15.82%	6.46%	3.10%	1.58%	0.78%	0.33%	**0.05%**	-0.13%	-0.24%
34	13.04%	5.67%	2.97%	1.72%	1.04%	0.65%	0.39%	0.22%	**0.11%**
35	10.68%	4.84%	2.66%	1.63%	1.06%	0.71%	0.49%	0.33%	**0.22%**
36	10.02%	5.30%	3.52%	2.67%	2.19%	1.90%	1.70%	1.56%	**1.46%**
37	7.82%	4.05%	2.63%	1.95%	1.57%	1.33%	1.16%	1.05%	**0.96%**
38	6.19%	3.15%	2.01%	1.47%	1.16%	0.97%	0.84%	0.75%	**0.68%**
39	4.84%	2.37%	1.46%	1.03%	0.79%	0.64%	0.54%	0.47%	**0.42%**
40	3.94%	1.91%	1.18%	0.84%	0.65%	0.54%	0.46%	0.41%	**0.37%**
41	3.06%	1.38%	0.79%	0.53%	0.38%	0.29%	0.24%	0.20%	**0.17%**
42	2.40%	1.02%	0.54%	0.33%	0.21%	0.15%	0.11%	0.08%	**0.05%**

Table A.70: *Estimated put prices as a percentage of market price for each fitted non-standardized Student distribution*

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8	η=9	η=10
Average	35.97%	13.13%	4.69%	1.33%	2.56%	3.62%	4.32%	4.81%	5.18%
St. Dev.	0.46	0.16	0.05	0.01	0.03	0.05	0.06	0.07	0.08
Maximum	165.52%	59.10%	19.24%	2.67%	10.86%	17.62%	22.19%	25.43%	27.86%
Minimum	2.40%	1.02%	0.54%	0.00%	0.11%	0.15%	0.05%	0.08%	0.05%

Table A.71: *Relative error statistics for estimated put prices for each fitted non-standardized Student distribution*

Statistic	η=2	η=3	η=4	η=5	η=6	η=7	η=8	η=9	η=10
Average	35.97%	13.13%	4.69%	0.72%	-1.47%	-2.79%	-3.66%	-4.25%	-4.68%
St. Dev.	0.46	0.16	0.05	0.01	0.04	0.05	0.07	0.07	0.08
Maximum	165.52%	59.10%	19.24%	2.67%	2.19%	1.90%	1.70%	1.56%	1.46%
Minimum	2.40%	1.02%	0.54%	-2.30%	-10.86%	-17.62%	-22.19%	-25.43%	-27.86%
Best dist.	0	0	0	6	2	0	1	0	9

Table A.72: *Absolute error statistics for estimated put prices for each fitted non-standardized Student distribution*

We have estimated out-of-sample option prices; all option prices verify no arbitrage conditions.

Strike k	Estimated $\hat{c}_k$								
	η=2	η=3	η=4	η=5	η=6	η=7	η=8	η=9	η=10
16	1.18%	**-0.37%**	-0.68%	-0.78%	-0.82%	-0.85%	-0.86%	-0.87%	-0.87%
18	1.71%	**-0.24%**	-0.66%	-0.81%	-0.88%	-0.92%	-0.94%	-0.96%	-0.97%
19	2.09%	**-0.13%**	-0.62%	-0.81%	-0.89%	-0.94%	-0.97%	-1.00%	-1.01%
20	2.57%	**0.04%**	-0.55%	-0.78%	-0.89%	-0.95%	-0.99%	-1.02%	-1.04%
37	331.24%	118.35%	59.24%	33.35%	19.29%	10.65%	4.88%	**0.76%**	-2.24%
38	437.00%	147.92%	69.33%	35.33%	16.92%	5.67%	**-1.92%**	-7.25%	-11.25%
39	617.13%	202.88%	93.63%	47.13%	22.38%	7.25%	**-2.75%**	-9.88%	-15.13%
40	760.33%	233.83%	99.67%	44.17%	15.00%	**-2.33%**	-14.00%	-22.00%	-28.00%
41	1072.00%	319.50%	135.25%	61.00%	23.25%	**0.75%**	-13.75%	-23.75%	-31.00%
42	1330.00%	374.33%	149.67%	62.33%	18.67%	**-6.33%**	-22.33%	-33.33%	-41.00%

Table A.73: *Estimated call prices for testing options as a percentage of market price*

Strike k	Estimated $\hat{p}_k$								
	η=2	η=3	η=4	η=5	η=6	η=7	η=8	η=9	η=10
16	699.00%	197.00%	42.00%	**-20.50%**	-50.00%	-66.00%	-75.50%	-81.50%	-85.00%
18	722.00%	231.20%	72.00%	**4.80%**	-29.20%	-48.00%	-59.60%	-67.60%	-72.80%
19	678.67%	228.33%	78.33%	**13.33%**	-20.00%	-39.33%	-51.33%	-59.67%	-65.33%
20	565.75%	194.75%	68.00%	**11.75%**	-17.75%	-35.25%	-46.50%	-54.25%	-59.75%
21	509.40%	184.20%	70.60%	19.00%	**-8.80%**	-25.60%	-36.60%	-44.20%	-49.80%
22	367.20%	130.67%	46.00%	**6.67%**	-14.80%	-28.00%	-36.80%	-43.07%	-47.60%
23	268.64%	93.36%	29.36%	**-0.91%**	-17.73%	-28.18%	-35.18%	-40.18%	-43.82%
24	215.40%	76.53%	25.00%	**0.27%**	-13.60%	-22.27%	-28.07%	-32.27%	-35.40%

Table A.74: *Estimated put prices for testing options as a percentage of market price*

Strike k	Estimated option price								
	η=2	η=3	η=4	η=5	η=6	η=7	η=8	η=9	η=10
$31 call	27.46%	5.34%	**-0.62%**	-3.03%	-4.20%	-4.84%	-5.20%	-5.42%	-5.56%
$32 call	40.13%	9.63%	**1.35%**	-2.01%	-3.66%	-4.56%	-5.09%	-5.42%	-5.62%
$31 put	24.83%	8.78%	3.15%	0.68%	**-0.57%**	-1.25%	-1.65%	-1.89%	-2.04%
$32 put	19.03%	6.98%	2.73%	0.84%	**-0.11%**	-0.64%	-0.96%	-1.15%	-1.27%
Average	27.86%	7.68%	1.65%	**-0.88%**	-2.14%	-2.82%	-3.23%	-3.47%	-3.62%
Abs. avrg.	27.86%	7.68%	1.96%	**1.64%**	2.14%	2.82%	3.23%	3.47%	3.62%

Table A.75: *Estimated option prices for ATM options as a percentage of market price*

Strike k	Estimated option price								
	η=2	η=3	η=4	η=5	η=6	η=7	η=8	η=9	η=10
$32	40.13%	9.63%	**1.35%**	-2.01%	-3.66%	-4.56%	-5.09%	-5.42%	-5.62%
$33	60.33%	16.95%	4.99%	**0.03%**	-2.46%	-3.88%	-4.73%	-5.28%	-5.65%
$34	91.74%	28.74%	11.10%	3.58%	**-0.32%**	-2.58%	-4.02%	-4.98%	-5.65%
$35	142.24%	48.44%	21.95%	10.44%	4.32%	**0.66%**	-1.73%	-3.37%	-4.54%
$36	219.09%	77.85%	38.00%	20.53%	11.09%	5.36%	1.55%	**-1.13%**	-3.06%
$37	331.24%	118.35%	59.24%	33.35%	19.29%	10.65%	4.88%	**0.76%**	-2.24%
$38	437.00%	147.92%	69.33%	35.33%	16.92%	5.67%	**-1.92%**	-7.25%	-11.25%
$39	617.13%	202.88%	93.63%	47.13%	22.38%	7.25%	**-2.75%**	-9.88%	-15.13%
$40	760.33%	233.83%	99.67%	44.17%	15.00%	**-2.33%**	-14.00%	-22.00%	-28.00%
$41	1072.00%	319.50%	135.25%	61.00%	23.25%	**0.75%**	-13.75%	-23.75%	-31.00%
$42	1330.00%	374.33%	149.67%	62.33%	18.67%	**-6.33%**	-22.33%	-33.33%	-41.00%
Average	463.75%	143.49%	62.20%	28.72%	11.32%	**0.97%**	-5.81%	-10.51%	-13.92%
Abs. avrg.	463.75%	143.49%	62.20%	29.08%	12.49%	**4.55%**	6.98%	10.65%	13.92%
Best dist.	0	0	1	1	1	**4**	2	2	0

Table A.76: *Estimated option prices for OTM call options as a percentage of market price*

Risk neutral default probabilities

Degrees of freedom	$\mathbb{P}(X \leq 0)$
2	0.4954%
3	0.1253%
4	0.0403%
5	0.0152%
6	0.0064%
7	0.0030%
8	0.0015%
9	0.0008%
10	0.0004%

Table A.78: *Three-month risk neutral probabilities of default for fitted non-standardized Student distributions*

Degrees of freedom	Annualized probabilities	Closest S&P rating
2	1.9669%	B+ (2.23%)
3	0.5003%	BB+ (0.40%)
4	0.1611%	BBB (0.19%)
5	0.0608%	A+ (0.06%)
6	0.0256%	AA- (0.03%)

7	0.0120%	AA (0.02%)
8	0.0060%	AA+, AAA (0.00%)
9	0.0032%	AA+, AAA (0.00%)
10	0.0016%	AA+, AAA (0.00%)

Table A.79: *One-year risk neutral probabilities of default for fitted non-standardized Student distributions assuming independence and closest matching S&P rating*

Degrees of freedom	Annualized probabilities	Closest S&P rating
2	1.6015%	BB- (1.09%)
3	0.7013%	BB (0.64%)
4	0.3797%	BB+ (0.40%)
5	0.2341%	BBB (0.19%)
6	0.1577%	BBB+ (0.13%)
7	0.1132%	BBB+ (0.13%)
8	0.0853%	A- (0.08%)
9	0.0667%	A (0.07%)
10	0.0538%	A+ (0.06%)

Table A.80: *One-year risk neutral probabilities of default for fitted non-standardized Student distributions assuming non independence and closest matching S&P rating*

In 2014, S&P had a LT local issuer rating of "A-" for Morgan Stanley. The Global Corporate Average Cumulative Default Rate for a one-year horizon corresponding to an "A-" rating is 0.08% – the probability has been estimated by S&P using data from 1981 to 2014.

Degrees of freedom	Annualized probabilities relative to S&P's one-year default probability
2	24.59
3	6.25
4	2.01
5	0.76
6	0.32
7	0.15
8	0.08
9	0.04
10	0.02

Table A.81: *Comparison of fitted default rates (independence hypothesis) and S&P default rates*

Degrees of freedom	Annualized probabilities relative to S&P's one-year default probability
2	20.02
3	8.77
4	4.75
5	2.93
6	1.97
7	1.42
8	1.07
9	0.83
10	0.67

Table A.82: *Comparison of fitted default rates (non independence hypothesis) and S&P default rates*

Risk neutral quantiles

Time to maturity for our panel of options is equal to 0.299, but based on a full year – i.e. taking into account weekend days. If we make the standard assumption that a year has 252 working days, multiplying this figure by 0.299 we obtain a time to maturity of 75 business days. Therefore to compute Morgan Stanley's quarterly stock returns we consider a 74-day lag:

$$R_t^{Trimestral} = \frac{S_t - S_{t-74}}{S_{t-74}}$$

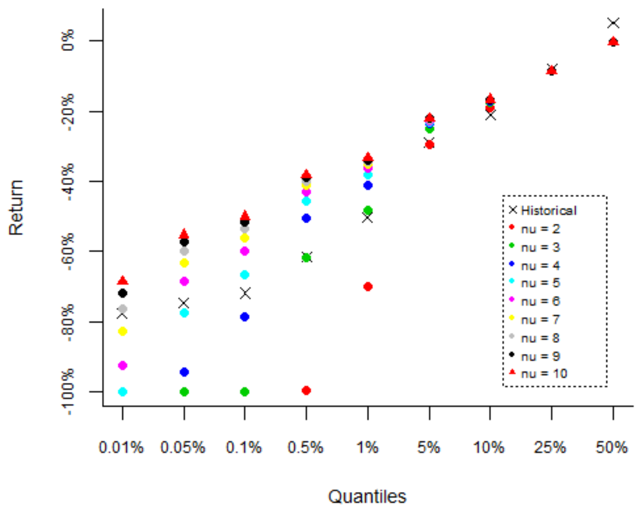
We have computed quantiles for the following stock price quarterly returns time series, ending on April 1st, 2014: 20 years (5,035 observations).

Quantile	25yr	η=2	η=3	η=4	η=5	η=6	η=7	η=8	η=9	η=10
0.01%	*-77.78%*	-100%	-100%	-100%	-100%	-92.38%	-82.61%	**-76.29%**	-71.91%	-68.70%
0.05%	*-74.70%*	-100%	-100%	-94.55%	**-77.55%**	-68.68%	-63.33%	-59.78%	-57.27%	-55.41%
0.1%	*-71.89%*	-100%	-100%	-78.82%	**-66.59%**	-60.07%	-56.08%	-53.41%	-51.51%	-50.09%
0.5%	*-61.56%*	-99.53%	**-61.88%**	-50.71%	-45.66%	-42.86%	-41.10%	-39.90%	-39.04%	-38.39%
1%	*-50.22%*	-69.95%	**-48.18%**	-41.34%	-38.16%	-36.38%	-35.26%	-34.49%	-33.94%	-33.52%
5%	*-29.07%*	**-29.52%**	-25.14%	-23.66%	-22.99%	-22.63%	-22.41%	-22.26%	-22.17%	-22.10%
10%	*-21.10%*	**-19.18%**	-17.60%	-17.11%	-16.93%	-16.85%	-16.82%	-16.81%	-16.81%	-16.81%
25%	*-7.97%*	**-8.50%**	-8.40%	-8.44%	-8.51%	-8.57%	-8.62%	-8.67%	-8.71%	-8.74%
50%	*5.12%*	-0.34%	-0.34%	-0.34%	-0.34%	-0.34%	-0.34%	-0.34%	-0.34%	-0.34%

Table A.83: *Relevant quantiles for fitted non-standardized Student distributions Together with historical quantiles for 20-year returns*

Degrees of freedom	Sum of squared errors
2	0.3789
3	0.1986
4	0.1207
5	0.1012
6	0.1020
7	0.1138
8	0.1378
9	0.1622
10	0.1843

Table A.84: *Sum of squared errors for each fitted non-standardized Student distribution*



Graph A.22: *Historical and distributional quantiles for selected distributions*

B.4/ Citigroup results with restricted choice of option prices

In this section, we have undertaken the same procedures as previously, but only keeping options for which the bid-ask spread is lower than or equal to 10%. Our goal is to measure the sensitivity of our procedure to the quantity and the quality of the data. In this section we have skipped the option pricing analysis and instead concentrated in probabilistic results.

B.4.1/ Data

Calls					Puts				
Strike	*Bid-offer spread*	*Mid*	*Volume*	*Implied volatility*	*Strike*	*Bid-offer spread*	*Mid*	*Volume*	*Implied volatility*
25.0	16.5%	21.60	-	53.08%	25.0	-	0.03	-	-
26.0	18.9%	20.58	-	-	26.0	-	0.04	-	48.49%
27.0	19.0%	19.55	-	-	27.0	-	0.04	-	45.64%
28.0	19.8%	18.58	-	-	28.0	-	0.04	-	42.88%
29.0	13.3%	17.70	-	50.08%	29.0	700.0%	0.05	-	41.63%
30.0	13.4%	16.70	-	46.94%	30.0	700.0%	0.05	-	38.97%
31.0	14.3%	15.70	-	43.89%	31.0	800.0%	0.05	-	36.95%
32.0	13.0%	14.75	-	43.47%	32.0	400.0%	0.06	-	35.37%
33.0	15.2%	13.78	-	41.53%	33.0	1000.0%	0.06	-	32.83%
34.0	17.4%	12.83	-	40.40%	34.0	500.0%	0.07	-	31.13%
35.0	17.1%	11.73	-	33.59%	35.0	175.0%	0.08	-	29.00%
36.0	20.4%	10.80	-	33.57%	36.0	114.3%	0.11	-	28.49%
37.0	20.8%	9.83	-	31.45%	37.0	20.0%	0.17	4	28.24%
38.0	12.3%	9.08	-	34.54%	38.0	31.6%	0.22	1	27.37%
39.0	10.5%	8.05	-	30.81%	39.0	10.7%	0.30	1	26.59%
40.0	**2.9%**	**7.05**	**-**	**27.64%**	**40.0**	**7.9%**	**0.40**	**1**	**25.87%**
41.0	**7.6%**	**6.13**	**-**	**25.77%**	**41.0**	**3.8%**	**0.53**	**16**	**25.27%**
42.0	**9.9%**	**5.30**	**-**	**25.13%**	**42.0**	**4.3%**	**0.71**	**19**	**24.71%**
43.0	**3.3%**	**4.58**	**-**	**25.22%**	**43.0**	**3.3%**	**0.93**	**34**	**24.16%**
44.0	**2.6%**	**3.85**	**-**	**24.60%**	**44.0**	**1.7%**	**1.20**	**16**	**23.64%**
45.0	**3.2%**	**3.20**	**82**	**24.19%**	**45.0**	**2.0%**	**1.55**	**73**	**23.25%**
46.0	**1.5%**	**2.61**	**-**	**23.74%**	**46.0**	**1.5%**	**1.96**	**23**	**22.84%**
47.0	**1.4%**	**2.10**	**44**	**23.38%**	**47.0**	**1.2%**	**2.45**	**92**	**22.54%**
48.0	**1.8%**	**1.66**	**49**	**23.09%**	**48.0**	**2.0%**	**3.02**	**8**	**22.40%**
49.0	**2.3%**	**1.30**	**90**	**22.95%**	**49.0**	**2.8%**	**3.65**	**38**	**22.11%**
50.0	**3.1%**	**1.00**	**164**	**22.79%**	**50.0**	**2.3%**	**4.35**	**6**	**21.88%**
52.5	**4.3%**	**0.48**	**22**	**22.47%**	**52.5**	**8.2%**	**6.35**	**-**	**21.43%**
55.0	**4.8%**	**0.22**	**81**	**22.35%**	55.0	12.0%	8.38	-	-
57.5	**10.0%**	**0.11**	**16**	**22.97%**	57.5	15.5%	10.78	-	-
60.0	100.0%	0.08	1	25.08%	60.0	23.3%	13.18	-	-
65.0	800.0%	0.05	-	29.52%	65.0	0.8%	18.43	-	-

***Tables A.85 and A.86**: Option prices data for Citigroup on April 7th, for a 103 days maturity (2014)*

B.4.2/ Risk neutral density estimation

The methodology implemented is the same. For comparison purposes, we have displayed results for our larger dataset.

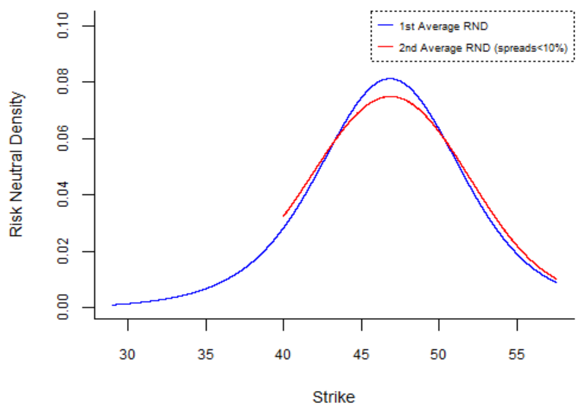
Method	Polynomial interpolation	Spline smoothing	Nadaraya-Watson estimator: h = h*	Nadaraya-Watson estimator: h = 9
SSE for implied volatility estimation	0.0000921	0.0001137	0.0001371	0.0007619
Results for larger dataset	0.0031028	0.0031403	0.0032261	0.0093162

Table A.87: *Sum of Squared Errors for all implied volatility estimation methods*

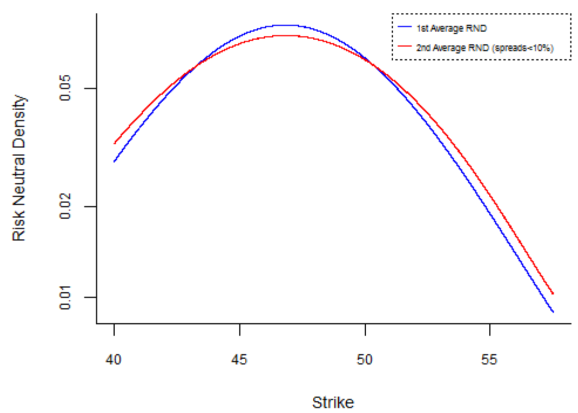
Method	Polynomial interpolation	Spline smoothing	Nadaraya-Watson estimator: h = 9
Density integral estimate	87.58%	85.45%	83.42%
Results for Larger dataset	95.42%	96.62%	94.64%

Table A.88: *Density integral estimates for polynomial, spline and Nadaraya-Watson methods*

Overall, the integral estimate is lower – around -8% – for our larger dataset. This is due to the fact that, our second dataset being more restricted, the density covers a smaller interval – \$40-\$57.5 against \$29-\$57.5 – thus the value of the integral for a smaller interval is mechanically lower than when the interval is larger. This might be considered a weakness of the density estimation method, as it might be the case – although we have not faced it for our datasets – than the estimate of the integral value is greater than 1. In such cases, adding further constraints might be necessary to avoid this problem; also, fitting a parametric distribution to the estimated RN distribution allows circumventing this issue.



Graph A.23: *Comparing average estimated RND for Citigroup*



Graph A.24: *Comparing average estimated RND for Citigroup in logarithmic scale*

Both distributions seem relatively close, although restricting our choice of options results in fatter tails.

B.4.3/ Student distribution fitting

Degrees of freedom	**σ** *(visual approximation)*	**σ** *(optimized fit)*	*Results for larger dataset*
2	\$4.70	\$4.9103	\$3.6681
3	\$4.90	\$5.0308	\$4.0148
4	\$5.10	\$5.1016	\$4.2669
5	\$5.15	\$5.1497	\$4.4547
6	\$5.20	\$5.1849	\$4.5990
7	\$5.225	\$5.2118	\$4.7134
8	\$5.225	\$5.2332	\$4.8058
9	\$5.225	\$5.2505	-
10	\$5.26	\$5.2649	-
11	\$5.28	\$5.2771	-
12	\$5.30	\$5.2874	-

Table A.89: *Optimal scale parameters for fitted non-standardized Student distributions*

Degrees of freedom	**Sum of squared errors**	*Results for larger dataset*	**Sum of log squared errors**	*Results for larger dataset*
2	7.37	11.8724	4,132.09	18,886.71
3	3.47	5.6819	1,877.60	8,001.47
4	2.05	2.7473	1,081.06	3,409.00
5	1.40	1.4162	716.27	1,459.44
6	1.06	0.8745	522.21	724.36
7	0.86	**0.7312**	408.42	**578.47**
8	0.73	0.7930	337.01	725.56
9	0.65	-	289.90	-
10	0.59	-	257.62	-
11	0.55	-	234.86	-
12	0.52	-	218.42	-

Table A.90: *Errors for each fitted non-standardized Student distribution*

Risk neutral default probabilities

Degrees of freedom	$\mathbb{P}(X \leq 0)$	*Results for larger dataset*
2	0.5468%	0.3074%
3	0.1334%	0.0688%
4	0.0400%	0.0200%
5	0.0138%	0.0069%
6	0.0053%	0.0027%
7	0.0022%	0.0012%
8	0.0010%	0.0005%
9	0.0005%	-
10	0.0002%	-
11	0.0001%	-
12	0.0001%	-

Table A.91: *Three-month risk neutral probabilities of defaultfor fitted non-standardized Student distributions*

Degrees of freedom	Annualized probabilities	Closest S&P rating
2	2.1693%	B+ (2.23%)
3	0.5325%	BB (0.64%)
4	0.1599%	BBB+ (0.13%)
5	0.0552%	A+ (0.06%)
6	0.0212%	AA (0.02%)
7	0.0088%	AA+, AAA (0.00%)
8	0.0040%	AA+, AAA (0.00%)
9	0.0020%	AA+, AAA (0.00%)
10	0.0008%	AA+, AAA (0.00%)
11	0.0004%	AA+, AAA (0.00%)
12	0.0004%	AA+, AAA (0.00%)

Table A.92: *One-year risk neutral probabilities of default for fitted non-standardized Student distributions assuming independence and closest matching S&P rating*

Degrees of freedom	Annualized probabilities	Closest S&P rating
2	1.8619%	B+ (2.23%)
3	0.8063%	BB (0.64%)
4	0.4178%	BB+ (0.40%)
5	0.2437%	BBB (0.19%)
6	0.1548%	BBB+ (0.13%)
7	0.1048%	A- (0.08%)
8	0.0746%	A (0.07%)
9	0.0552%	A+ (0.06%)
10	0.0423%	AA- (0.03%)
11	0.0332%	AA- (0.03%)
12	0.0267%	AA- (0.03%)

Table A.93: *One-year risk neutral probabilities of default for fitted non-standardized Student distributions assuming non independence and closest matching S&P rating*

We remind the reader that Citigroup rating was “A” on April 2014, corresponding to a historical default rate of 0.07%.

Degrees of freedom	Annualized probabilities relative to S&P’s one-year default probability (independence hyp.)	Annualized probabilities relative to S&P’s one-year default probability (non-independence hyp.)
2	6.84	27.12
3	1.67	6.66
4	0.50	2.00
5	0.17	0.69
6	0.07	0.27
7	0.03	0.11
8	0.01	0.05
9	0.01	0.03
10	0.00	0.01
11	0.00	0.01
12	0.00	0.01

Table A.94: *Comparison of fitted default rates (independence hypothesis in the middle column and non-independence hypothesis in the right column) and S&P default rates*

Risk neutral quantiles

Quantile	30yr	η=2	η=3	η=4	η=5	η=6	η=7	η=8	η=9	η=10	η=11	η=12
0.01%	*-85.6%*	-100%	-100%	-100%	-100%	**-89.3%**	-79.0%	-72.4%	-67.8%	-64.4%	-61.8%	-59.7%
0.05%	*-82.6%*	-100%	-100%	-94.3%	**-76.0%**	-66.3%	-60.5%	-56.6%	-53.9%	-51.8%	-50.3%	-49.0%
0.1%	*-80.6%*	-100%	-100%	**-78.6%**	-65.2%	-58.0%	-53.5%	-50.6%	-48.4%	-46.8%	-45.6%	-44.6%
0.5%	*-72.7%*	-100%	**-63.1%**	-50.4%	-44.6%	-41.3%	-39.1%	-37.7%	-36.6%	-35.8%	-35.2%	-34.7%
1%	*-59.7%*	-73.4%	**-49.0%**	-41.0%	-37.2%	-35.0%	-33.5%	-32.5%	-31.8%	-31.2%	-30.8%	-30.4%
5%	*-31.1%*	**-30.8%**	-25.4%	-23.3%	-22.3%	-21.6%	-21.2%	-20.9%	-20.6%	-20.5%	-20.3%	-20.2%
10%	*-19.6%*	**-19.9%**	-17.7%	-16.8%	-16.3%	-16.0%	-15.8%	-15.7%	-15.6%	-15.5%	-15.4%	-15.4%
25%	*-5.9%*	-8.6%	-8.2%	-8.1%	-8.0%	-8.0%	-7.9%	-7.9%	-7.9%	-7.9%	-7.9%	**-7.9%**
50%	*4.8%*	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%	0.04%

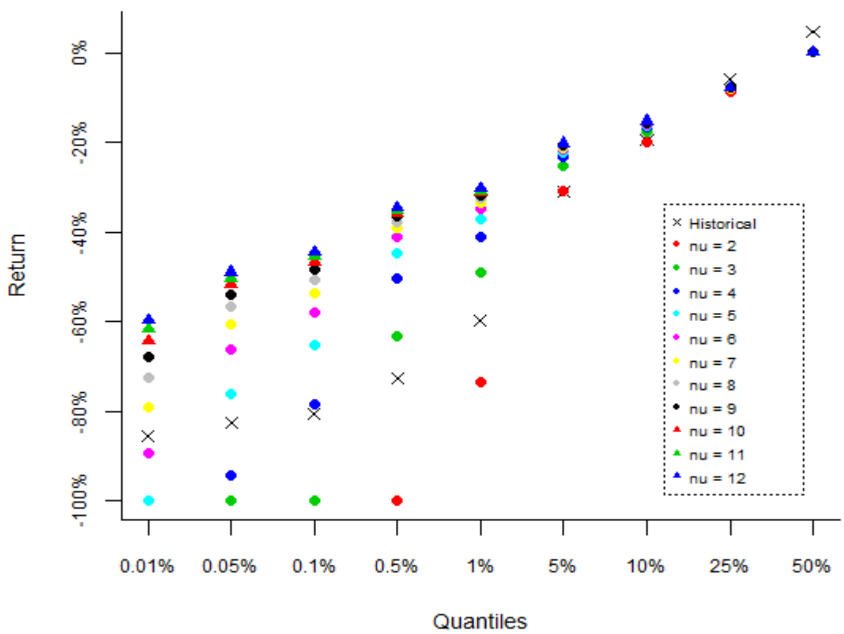
Table A.96: *Relevant quantiles for fitted non-standardized Student distributions together with historical quantiles for 30-year returns*

Degrees of freedom	Sum of squared errors	*Results for larger dataset*
2	0.1852	0.1054
3	0.1159	0.1660
4	0.1289	0.2205
5	0.1900	0.3036
6	0.2517	0.3775
7	0.3209	0.4458
8	0.3858	0.5000
9	0.4406	-
10	0.4861	-
11	0.5239	-
12	0.5557	-

Table A.97: *Sum of squared errors for each fitted non-standardized Student distribution*

Dataset	30yr	η=2	η=3	η=4	η=5	η=6	η=7	η=8	η=9	η=10	η=11	η=12
Large	-59.71%	**-54.84%**	-39.12%	-34.31%	-32.16%	-31.01%	-30.32%	-29.86%	-	-	-	
Restricted	-59.71%	-73.4%	**-49.0%**	-41.0%	-37.2%	-35.0%	-33.5%	-32.5%	-31.8%	-31.2%	-30.8%	-30.4%

Table A.98: *Comparison of values for the 1% quantile for two different datasets*



Graph A.25*: Historical and distributional quantiles for selected distributions*

PARIS

ESSEC Business School
3, avenue Bernard-Hirsch
CS 50105 Cergy
95021 Cergy-Pontoise Cedex
France
Tél. +33 (0)1 34 43 30 00
www.essec.edu

ESSEC Executive Education
CNIT BP 230
92053 Paris-La Défense
France
Tél. +33 (0)1 46 92 49 00
www.executive-education.essec.fr

ESSEC Asia-Pacific
5 Nepal Park
Singapore 139408
Tél. +65 6884 9780
www.essec.edu/asia

SINGAPOUR

Contact :

Centre de Recherche
+33 (0)1 34 43 30 91
research.center@essec.fr

ISSN 1291-9616

affilié à la
CCI PARIS ILE-DE-FRANCE

ESSEC BUSINESS SCHOOL